

FICHE 1

DÉFINITIONS ET USAGES DE L'IA EN SANTÉ

Martin Cousineau, Alexandre Castonguay,
Philippe Després, Aude Motulsky, Jean Noel Nikiema,
Janine Badr, Joé T. Martineau, Cécile Petitgand,
Catherine Régis

Québec, Canada
Mars 2022

Mandat réalisé par un groupe de travail de l'Axe Santé durable de l'OBVIA pour la Direction de la planification stratégique et de la performance (DGPPS) du MSSS



**OBSERVATOIRE INTERNATIONAL
SUR LES IMPACTS SOCIAUX
DE L'IA ET DU NUMÉRIQUE**



**OBSERVATOIRE INTERNATIONAL
SUR LES IMPACTS SOCIÉTAUX
DE L'IA ET DU NUMÉRIQUE**

L'OBVIA est un réseau de recherche ouvert qui contribue aux réflexions entourant l'utilisation responsable de l'IA et du numérique en favorisant les discussions et le maillage entre les communautés de pratique, la société civile et les décideurs publics. Soutenu financièrement par les Fonds de recherche du Québec (FRQ), il fédère les expertises de plus de 260 chercheur-es couvrant les secteurs des sciences humaines et sociales, des sciences et génie et de la santé. Il bénéficie également de l'appui de plus de 125 partenaires issus de centres de recherche et d'organisations publiques et privées du Québec, du Canada et de l'étranger.

Son objectif est de mettre en lumière les enjeux cruciaux soulevés par les développements et l'utilisation de l'IA et du numérique et d'identifier des solutions plausibles aux problèmes et opportunités qu'ils soulèvent.

L'OBVIA réalise des travaux de recherche et des mandats d'intervention en partenariat en vue de développer de nouvelles connaissances, de les diffuser et d'aider les organisations québécoises à maximiser les impacts positifs du numérique et de l'IA dans nos sociétés. Ces travaux sont priorisés en fonction des orientations scientifiques de l'OBVIA et visent à répondre à la fois aux objectifs des partenaires et des équipes de recherche impliquées. Ils peuvent prendre plusieurs formes, comme des études sur le terrain, des synthèses et analyses de la littérature, la conception d'ateliers ainsi que le soutien dans l'élaboration, l'adoption et la mise en œuvre de bonnes pratiques en matière d'enjeux touchants à l'IA et au numérique.

**Observatoire international sur les
impacts sociétaux de l'intelligence artificielle et du numérique**

Pavillon Charles-De Koninck, local 2489
1030, avenue des Sciences-Humaines
Université Laval
Québec (Québec) G1V 0A6

observatoireia@ulaval.ca
observatoire-ia.ulaval.ca

ISBN : 978-2-925138-14-3

Liste des abréviations

ACMTS	Agence canadienne des médicaments et des technologies de la santé
ANS	Agence nationale de santé (France)
BCH	Bureau de la certification et de l'homologation
C.c.Q.	Code civil du Québec
CDPJ	Commission des droits de la personne et des droits de la jeunesse
CHUM	Centre hospitalier de l'Université de Montréal
CNN	Conseil canadien des normes
DIM	Direction des instruments médicaux
DRI	Direction des ressources informationnelles
DSQ	Dossier santé Québec
ECG	Électrocardiographie
EEG	Électroencéphalographie
ETS	Évaluation des technologies de santé
FDA	Food and Drug Administration
GDHP	Global Digital Health Partnership
GPUs	Graphics Processing Units (Unités de traitement graphique)
IA	Intelligence artificielle
INR	Institut national de recherche
ISO	Organisation internationale de normalisation
LRDPDE	Loi sur la protection des renseignements personnels et des documents électroniques
MHRA	Medicines and Healthcare Products Regulatory Agency
NASA	National Aeronautics and Space Administration
NHS	National Health System (Royaume Uni)
NLP	Natural Language Processing (Traitement automatique de la langue)
OCDE	Organisation de coopération et de développement économique
OMS	Organisation mondiale de la santé
QALYs	Quality-adjusted life years (Année de vie ajustée pour la qualité)
SIA	Systèmes d'intelligence artificielle
TGV	Trousse globale de vérification
TRL	Technology Readiness Level (Niveau de maturité technologique)
UE	Union Européenne
VIH	Virus de l'immunodéficience humaine

Table des matières

1. Contexte	5
2. Démarche	6
3. Les sens de l'IA	7
4. Classification des systèmes d'IA	9
4.1. Contexte.....	9
4.2. Intrants ou données.....	10
4.3. Modèle d'IA.....	11
4.4. Extrants ou tâches.....	12
5. Recommandations entourant l'adoption d'une définition utile de l'IA en santé et autres pistes de réflexion	14
6. En savoir plus	16
Références	17

1 CONTEXTE

Plusieurs avancées technologiques majeures en lien avec les données numériques, la robotisation et les théories de l'intelligence artificielle (IA) ont vu le jour au cours des années 1950 et 1960, constituant ainsi les fondements de ce qui est aujourd'hui reconnu comme étant le domaine de l'IA (Russell & Norvig, 2021). Malgré le développement et les premiers cas d'adoption de plusieurs technologies numériques en santé, telles que les dossiers médicaux électroniques dans les années 1960 et les systèmes d'aide au diagnostic et à la prise en charge des patients (c.-à-d., des systèmes experts) dans les années 1970, les capacités technologiques de l'époque ont grandement limité l'intérêt envers les **systèmes d'IA (SIA)**. Durant les années 1990, de nouvelles avancées théoriques et une capacité numérique accrue ont permis à l'IA de réellement remplir ses promesses (Kaul et al., 2020). Depuis les années 2010, un regain d'intérêt pour l'IA est constaté tant en recherche et développement, qu'en termes d'investissements publics et privés. Cet intérêt renouvelé, soutenu par un impressionnant développement logiciel et matériel, ont amené l'IA à un niveau de maturité permettant d'accomplir des tâches de plus en plus complexes (AI4H, 2020). Ceci est en bonne partie dû au volume de données disponibles sous forme numérique dans plusieurs sphères d'activité humaine, à la puissance de calcul (p. ex., les unités de traitement graphique - *Graphics Processing Units* ou *GPUs*) et aux avancées scientifiques (p. ex., les réseaux profonds de neurones artificiels). Aujourd'hui, le niveau de maturité de l'IA est démontré par plusieurs cas d'utilisation réels; plusieurs SIA sont disponibles et prêts à être déployés dans un grand nombre de secteurs d'activité, incluant celui de la santé. Il devient primordial pour les acteurs de ce secteur de mieux comprendre ce qu'est l'IA, ce qu'elle permet - ou pas - de faire (fiche 2), et ce qu'elle soulève comme enjeux légaux (fiche 3) et éthiques (fiche 4). Un premier pas consiste à définir ce qu'est l'IA, ce que cette fiche propose de faire par l'entremise d'une vision pratique et utile de cette technologie, dans le contexte de définitions mouvantes.



2 DÉMARCHE

Une revue de la littérature scientifique et grise portant spécifiquement sur les SIA en santé a été menée, en portant une attention particulière aux initiatives d'organismes à portée internationale tels que l'Union européenne (UE) et l'Organisation de coopération et de développement économiques (OCDE). Deux documents de ces organisations, quoique généraux, ont grandement contribué au contenu de cette fiche. Le rapport de l'UE, préparé par AI Watch¹ en s'appuyant sur une méthodologie rigoureuse, présente une synthèse des définitions de l'IA les plus largement adoptées à travers le monde (Samoili et al., 2021). Ce document présente également une taxonomie, une définition englobante de l'IA, ainsi que des pistes de réflexion entourant l'adoption d'une définition utile. Un second document, de l'OCDE, rapporte l'élaboration d'un système de classification des SIA. Ce cadre s'avère très utile à une meilleure compréhension des caractéristiques d'un SIA. D'autres références particulièrement pertinentes aux définitions de l'IA en santé sont rapportées dans la fiche; la littérature associée est abondante et croît à un rythme soutenu. Finalement, cette fiche a fait l'objet d'un atelier délibératif afin d'identifier les principales tâches qui peuvent être accomplies par des SIA en santé.

¹ Le service de connaissance de la Commission [européenne] chargé de surveiller le développement, l'adoption et l'impact de l'intelligence artificielle pour l'Europe.

3 LES SENS DE L'IA

Définir l'IA est un exercice périlleux, car cette expression peut avoir plusieurs sens. L'IA est d'abord une discipline, à l'interface de l'informatique, des mathématiques, des statistiques, du génie, voire des neurosciences, avec des approches théoriques, des méthodes, et des domaines d'étude et d'application qui lui sont propres. Les principaux domaines de l'IA, tels que définis par une taxonomie publiée par l'UE (Samoili et al., 2021), sont rapportés au Tableau 1. Une liste des principaux mots-clés associés à chaque sous-domaine de l'IA est également disponible (voir la section En savoir plus).

Tableau 1. Une taxonomie de l'IA

	Domaines de l'IA	Sous-domaines de l'IA
Domaines centraux	Raisonnement	Représentation des connaissances
		Raisonnement automatique
		Raisonnement de sens commun
	Planification	Planification et ordonnancement
		Recherche
		Optimisation
	Apprentissage	Apprentissage automatique
	Communication	Traitement naturel du langage
	Perception	Vision par ordinateur
		Traitement audio
Domaines transversaux	Intégration et interaction	Systemes multi-agents
		Robotique et automatisation
		Véhicules connectés et autonomes
	Services	Services de l'IA
	Éthique et philosophie	Éthique de l'IA
		Philosophie de l'IA

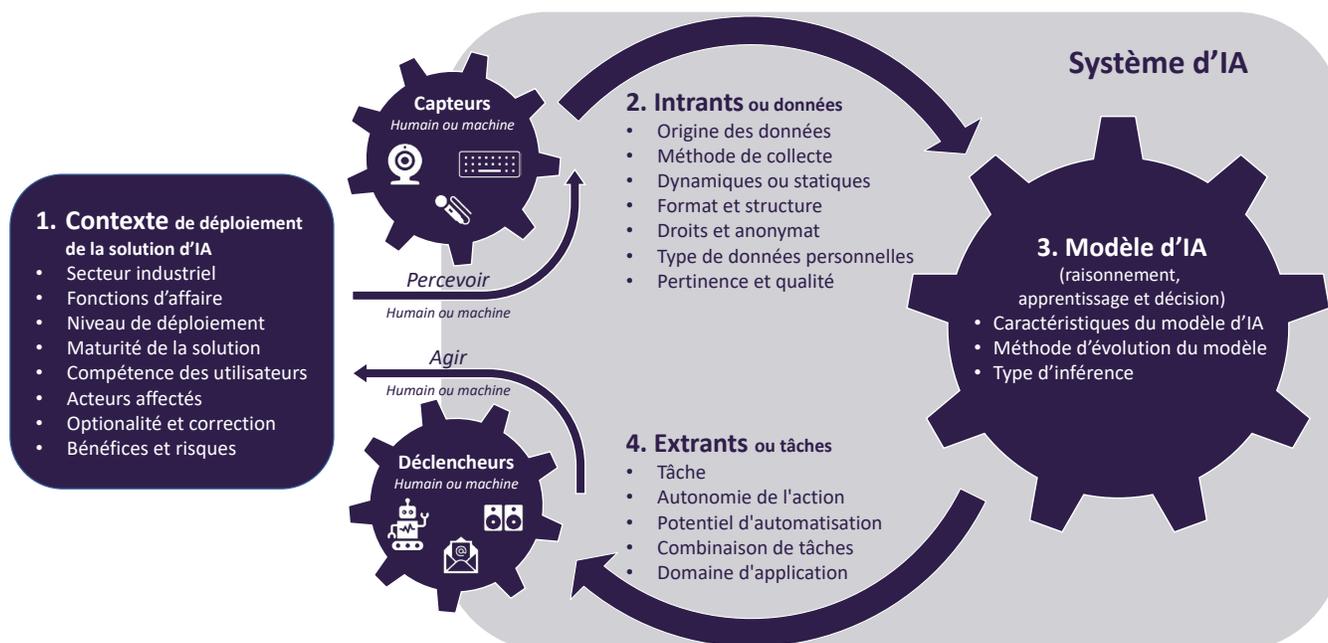
L'expression « intelligence artificielle » peut également être utilisée pour référer à un type de système possédant certaines habiletés et composantes, qui lui permettent d'effectuer des tâches cognitives qui sont fortement associées à l'intelligence humaine. En utilisant une liste de critères, une définition plus pragmatique des SIA est alors obtenue.

Un SIA est composé d'un modèle d'IA, qui utilise et dépend de capteurs et de déclencheurs pour interagir avec son contexte. Alors que les capteurs permettent au système de percevoir son environnement, les déclencheurs lui offrent la possibilité de prendre action et, possiblement, de modifier son environnement. La Figure 1 rapporte l'ensemble des critères de classification des SIA selon le cadre de l'OCDE (2022), tout en les présentant de sorte à être compatibles avec la vision conceptuelle de l'Union européenne (GEHN IA, 2018). Ce schéma, tout comme les cadres de référence qu'il représente, se veut le plus général possible mais est appelé à changer étant donné la célérité avec laquelle le domaine évolue.

Boîte 1. Modèle d'IA, système d'IA ou outil numérique?

Le **modèle d'IA** est l'algorithme en tant que tel. C'est l'élément que l'on cherche à définir ou à classifier selon les références présentées dans cette fiche. Un **système d'IA (SIA)** est composé d'un ou plusieurs modèles d'IA, mais aussi de capteurs, d'intrants (*inputs*), d'extrants (*outputs*) et de déclencheurs (voir Figure 1). C'est donc un système ayant plusieurs composantes interagissant avec un ou plusieurs modèles d'IA. Le ou les modèles d'IA œuvrent au sein d'un **outil numérique** (p. ex., logiciel, robot), lequel pourra faciliter l'interaction entre l'utilisateur et les modèles, mais aussi offrir d'autres fonctionnalités qui peuvent être liées ou non au domaine de l'IA *stricto sensu*.

Figure 1. Schéma des quatre dimensions d'un système d'IA



4 CLASSIFICATION DES SYSTÈMES D'IA

Pour certains, une définition universelle des SIA est un objectif inatteignable et potentiellement nuisible (Aiken, 2021). Sans être incompatible avec l'adoption d'une définition englobante des SIA sous toutes leurs formes, une approche davantage pragmatique que philosophique - c'est-à-dire de définir par la classification - est suggérée par l'OCDE (2022). Ce cadre international de classification des SIA facilite l'appréciation des impacts potentiels d'un SIA et le niveau de risque associé. Suivant ce cadre, chaque SIA présente quatre dimensions : contexte, données ou intrants, modèle d'IA, et tâches ou extrants (comme illustré à la Figure 1).

Afin de rendre cette fiche concise et accessible sans sacrifier les éléments les plus centraux, une sélection de sous-dimensions a été effectuée en s'inspirant de deux présentations abrégées préparées par l'OCDE (Addo et al., 2021; Murdick, 2020). Une priorité a été accordée aux critères de classification non-optionnels tel que définis par l'OCDE.

Cette classification ne force pas la sélection d'une caractéristique ou d'une autre. Il s'agit plutôt de sélectionner toutes les caractéristiques présentes dans le SIA à classifier. En effet, un même système d'IA pourrait traiter plusieurs types d'intrants en combinant différents types de modèles d'IA, pour finalement produire une variété d'extrants.

4.1. Contexte

Le contexte réfère à l'environnement socioéconomique, culturel, géographique, politique, etc. dans lequel est déployé le SIA, ce qui inclut, mais dépasse largement le lieu physique d'implantation. Voir ci-dessous, ainsi que la Fiche 2 pour plus de détails au sujet du contexte.

Secteur industriel et fonction d'affaires : Le **secteur industriel** (*industrial sector*) représente la catégorie d'activités économiques de haut niveau au sein duquel il est prévu que le SIA opère (p. ex., santé, éducation, transport et entreposage). La **fonction d'affaires** (*business function*) représente quant à elle le sous-secteur visé par le SIA, au sein d'une organisation (p. ex., gestion des ressources humaines, comptabilité, services cliniques).

Niveau de déploiement et maturité technologique : Les projets d'envergure sont souvent mis à l'essai par l'entremise de projets pilotes, avant de passer au travers de divers stades de **déploiement**, afin de contrôler le niveau de risque associé à l'implantation d'une nouvelle initiative. Les risques associés au déploiement ne sont pas les mêmes en fonction du niveau de **maturité** technologique (*technology readiness level - TRL*). À cet effet, des normes telles qu'ISO 16290 (ISO, 2013) sont fréquemment utilisées par des organisations publiques, parapubliques et privées telles que l'OCDE (2022), la NASA (Mankins, 1995), Innovation Canada (2018) ou MEDTEQ+ (2021).

Utilisateurs et acteurs affectés : Les **utilisateurs** d'un SIA diffèrent souvent des développeurs ou des personnes ayant pris en charge l'implantation. Dans le milieu de la santé, l'utilisateur pourrait être un clinicien ou un administrateur n'ayant aucune expérience en programmation. Il importe donc de déterminer le niveau d'expertise attendu (amateur, formé ou expert) par les utilisateurs d'un SIA, information qui pourra avoir des répercussions importantes au niveau du développement, par exemple. En plus de l'utilisateur, différents **acteurs** peuvent être affectés par un SIA. Par exemple, la patientèle d'un médecin est aussi affectée par le SIA que ce médecin utilise pour guider ses décisions cliniques.

4.2. Intrants ou données

Bien que plusieurs facteurs puissent affecter la qualité des données, celle-ci sera grandement tributaire du processus lié à leur génération, que ce soit dans un contexte contrôlé (p. ex., étude clinique) ou dans un contexte de routine (p. ex., activités cliniques aux soins intensifs, facturation des actes médicaux). La conception du SIA et la terminologie ou le codage utilisé vont contraindre la façon dont l'information sera colligée et organisée; par exemple, si un diagnostic n'est pas codé de la même façon d'un hôpital à l'autre, il sera complexe d'utiliser un même modèle pour ces différents hôpitaux. À ce titre, il est pertinent de rappeler que les normes liées aux données du domaine de la santé sont administrées par Info-route Santé du Canada. La méthode de collecte et de saisie, ainsi que la nature des intrants (c.-à-d., les données) qui sont utilisés par le SIA sont présentées ci-après.

Agent de collection : Deux types d'agents pourraient être appelés à colliger des données au sein d'un système d'information, soit un **humain**, manuellement, un **capteur**, de manière automatisée, ou une combinaison des deux. Par exemple, la tension artérielle d'une personne peut être captée de manière automatisée par un capteur et transmise automatiquement dans un SIA, ou alors mesurée par l'humain et saisie manuellement dans un SIA, ou finalement captée automatiquement puis saisie manuellement.

Provenance : Les données peuvent provenir d'**experts** (c.-à-d., des connaissances humaines saisies sous forme de règles ou de structures, telles que des ontologies ou des arbres décisionnels utilisés par l'IA symbolique), d'**actions** (p. ex., heure à laquelle un acte médical a été dispensé), d'**observations** (p. ex., signes vitaux), de résultats de **simulations** (p. ex., les réactions physiologiques estimées par un simulateur de chirurgie), et d'une **agrégation** de données (p. ex., groupes d'âge ou score d'Apgar).

Nature dynamique et structure : Les données peuvent être **dynamiques** (p. ex., mises à jour au fil du temps) ou **statiques** (p. ex., le code génétique d'une personne qui ne change pas). De plus, elles peuvent être **structurées** (p. ex., bases de données relationnelles avec des champs simples), **semi-structurées** (p. ex., radiographies accompagnées de métadonnées) ou **non structurées** (p. ex., texte, images, vidéos). Un type de structure se distingue de la compréhension commune du terme « donnée ». Il s'agit des **données complexes structurées**. Sont incluses dans cette catégorie les ontologies, les bases de connaissances, les règles (tel qu'on retrouve dans les systèmes experts) et les fonctions analytiques (p. ex., les fonctions d'apprentissage).²

Droits d'utilisation : Trois types de droits peuvent être associés aux données, qui parfois se chevauchent (voir la Fiche 3 pour plus de détails) : **propriétaires** (p. ex., droits d'auteurs et brevets), **publics** (aucun droit contraignant leur utilisation), **personnelles** (données concernant une personne). En contexte d'utilisation de données personnelles, il importe de sélectionner judicieusement le niveau d'anonymat approprié afin de protéger l'identité de la personne sans trop nuire à l'efficacité des analyses. Un exemple d'équilibre pourrait être l'utilisation d'un pseudonyme et le retrait des informations permettant une identification indirecte (p. ex., code postal, lieu de travail, poste occupé), mais la conservation des diagnostics reçus, de l'âge et du sexe.

² Ainsi, à l'instar du cadre de classification de l'OCDE qui catégorise ces types d'intrants comme un type de données, il est considéré dans le cadre de cette fiche que tous les intrants sont également des données.

4.3. Modèle d'IA

Le modèle d'IA réfère à l'algorithme qui typiquement permet de passer des intrants aux extrants. En IA, cet algorithme n'est pas nécessairement statique et peut *apprendre*, au sens machine, à partir de nouveaux intrants (de nouvelles données). Les extrants peuvent prendre la forme d'une description, d'une prédiction ou d'une prescription (c.-à-d., d'une décision).

Type statistique

On retrouve ici les modèles basés sur l'**apprentissage automatique** (*machine learning*) qui sont souvent au cœur des SIA contemporains. L'entraînement d'un modèle d'IA de type statistique peut se faire par :

- **apprentissage supervisé** (*supervised learning*) : de type « top-down », où le modèle tente d'exécuter une tâche donnée tout en suivant un plan prédéterminé. Dans cette approche, le modèle s'entraîne à partir de données pré-étiquetées; on présente au modèle une entrée et la sortie associée, qui a valeur de vérité (*ground truth*). Par exemple, un modèle d'IA analyse des radiographies déjà catégorisées, soit celles qui présentent une pathologie et celles qui n'en présentent pas. Cette catégorisation (ou étiquetage avec la vérité) a été faite par des experts au préalable. Le modèle d'IA entraîné avec cette banque de radiographies déjà correctement catégorisées tente par la suite de déterminer la présence de pathologies sur de nouvelles radiographies non-étiquetées. Ce type de modèle pourra ensuite être utilisé sur un jeu de données distinct de celui qui a été utilisé pour l'entraînement afin d'évaluer sa performance et éventuellement la comparer à celle d'agents humains. Certains défis sont associés à ce type d'approche, notamment en ce qui a trait à la capacité du modèle d'IA à bien performer au sein d'environnements différents (p. ex., différents hôpitaux) et auprès de populations variées (p. ex., s'assurer de pouvoir démontrer la performance du modèle d'IA pour différents sexes, groupes d'âge et ethnicités).
- **apprentissage non supervisé** (*unsupervised learning*) : de type « bottom-up », où le modèle tente d'exécuter une tâche donnée sans plan prédéterminé; le modèle construit son propre plan. Dans ce type d'approche, les données ne sont pas pré-étiquetées au préalable. Par exemple, le modèle reçoit un grand nombre de radiographies et est laissé libre de regrouper celles qui sont similaires afin d'identifier des catégories de pathologies. C'est un humain qui doit fournir une fonction qui permet de mesurer les similarités entre les radiographies, mais les catégories sont identifiées par le modèle. Un autre exemple important d'apprentissage non supervisé est les modèles récents de traitement automatique du langage naturel (*natural language processing - NLP*), qui réussissent à découvrir le sens (p. ex., les sentiments) sous-jacent au texte.
- **apprentissage par renforcement** (*reinforcement learning*) : il s'agit ici pour le modèle d'améliorer sa performance par essai-erreur. Prenons l'exemple d'un modèle d'IA déterminant les mouvements d'un robot de livraison autonome chargé de transporter des prélèvements entre le laboratoire d'analyses et les différents services cliniques d'un centre hospitalier. Ce modèle d'IA devra pouvoir s'adapter et trouver des solutions alternatives chaque fois qu'un obstacle se présentera le long de l'itinéraire prévu pour le robot. Le modèle d'IA devra adapter les mouvements du robot afin qu'il puisse s'engager sur des trajectoires alternatives qui le mèneront tout de même vers la destination souhaitée.

Type symbolique

Dans cette catégorie, l'algorithme du modèle d'IA ne permet aucune forme d'apprentissage automatique. L'algorithme, les variables, ainsi que les règles et leurs paramètres sont explicitement définis par l'humain en fonction d'une représentation humaine de la situation. Ces modèles sont construits sous forme d'arguments de type « si-alors » ou de formules mathématiques. On y retrouve, entre autres, les ontologies (c.-à-d., une structure organisant les concepts d'un domaine), les systèmes experts (ceux qui utilisent des règles) et les algorithmes de recherche tel que l'optimisation mathématique qui est largement utilisée dans le domaine de la recherche opérationnelle. Plusieurs de ces modèles font partie des premiers SIA et sont toujours d'actualité. Par exemple, les systèmes experts basés sur des règles étaient répandus dans les années 1980 et sont encore largement utilisés dans le secteur de la santé (voir Fiche 2). De même, l'optimisation mathématique, en particulier la programmation linéaire datant des années 1950 est encore utilisée afin d'améliorer l'efficacité des opérations tout en respectant diverses contraintes.

Type hybride

Plusieurs modèles d'IA sont souvent appelés à interagir les uns avec les autres, que ce soit au sein d'un même SIA ou avec des SIA distincts. Ces différents modèles pourront être de type statistique ou symbolique ou même combiner les deux. Par exemple, les algorithmes de traitement naturel du langage (NLP) combinent souvent des approches statistiques s'appuyant sur des données massives à des approches symboliques prenant en compte des structures telles que les règles de grammaire.

Boîte 2. Note sur l'interprétabilité et l'explicabilité

En général, plus un modèle d'IA est complexe, moins il sera interprétable par des utilisateurs qui tenteraient de comprendre comment un SIA est arrivé à un résultat donné. Bien que potentiellement complexes, les modèles d'IA symbolique (p. ex., les systèmes experts) sont construits à partir de représentations humaines et sont donc généralement plus simples à comprendre qu'un modèle d'IA statistique, car il est toujours possible de revenir aux règles pour expliquer le résultat. Les modèles basés sur l'apprentissage profond (*deep learning*), qui sont utilisés autant pour l'apprentissage supervisé, non supervisé et par renforcement, utilisent des réseaux de neurones artificiels qui sont réputés peu interprétables par les utilisateurs. L'explicabilité est un concept similaire qui réfère à la capacité des experts en IA d'expliquer comment, d'un point de vue algorithmique, le modèle arrive à produire un résultat donné.

4.4. Extrants ou tâches

Les extrants ou les tâches représentent ce qu'un SIA fait avec le résultat de ses analyses. Plus spécifiquement, un SIA pourrait avoir comme tâche d'effectuer des prédictions. L'extrant serait alors, par exemple, la production d'un score. Ensuite, un déclencheur, lequel peut être un humain ou une machine, peut alors agir en réaction à l'extrant, par exemple en activant une alerte, et ainsi potentiellement modifier l'environnement au sein duquel le système d'IA évolue. Le Tableau 2 contextualise les principales tâches d'un SIA en offrant des exemples avec différents types d'utilisateurs dans le secteur de la santé.

Tableau 2. Types de tâches d'un système d'IA en santé

Tâches	Explication	Exemples d'utilisation en santé
Reconnaissance (utilise surtout la classification supervisée)	Identifie et catégorise les données (p. ex., image, vidéo, audio et texte) en classes spécifiques (p. ex., « ceci est une tumeur »).	Patients : Un SIA utilisant les expressions faciales de l'utilisateur pour potentiellement détecter de l'insatisfaction, puis personnaliser les paramètres de l'intervention. Cliniciens : Classification des images de rétines pour soutenir le diagnostic du diabète. Gestionnaires : Transcription automatique de l'écriture manuscrite en texte numérique par la reconnaissance optique de caractères.
Détection d'événement (utilise principalement de l'apprentissage non supervisé et par renforcement)	Connecte des données entre elles en temps réel afin de détecter des valeurs aberrantes ou des anomalies (événements rares).	Patients : Détection de glycémie anormale selon des cibles prévues. Cliniciens : Aide à la décision clinique, p. ex., vérifier la présence de contre-indication dans une prescription considérant le dossier médical du patient. Gestionnaires : Soutien à la détection de parcours de soins anormaux (p. ex., très longs).
Prévision (utilise de l'apprentissage supervisé et des méthodes de prévision)	Utilise les événements passés et présents pour prédire des résultats futurs (événements fréquents, qui risquent de se reproduire plus tard).	Patients : Prédire le temps d'attente à l'urgence à un moment donné. Cliniciens : Déterminer le risque de réadmission à l'hôpital en fonction de l'évolution de certains indicateurs cliniques. Gestionnaires : Prédire quel volume d'équipements de protection individuelle est nécessaire pour la prochaine semaine.
Personnalisation (utilise de l'apprentissage supervisé et par renforcement)	Développe un profil pour une personne, puis tente de s'adapter à cette personne au fil du temps.	Patients : Ajuster la dose d'insuline selon l'alimentation et l'historique du patient. Cliniciens : Aide à la décision clinique tenant compte des caractéristiques du patient pour soutenir le meilleur choix de traitement. Gestionnaires : Soutien à la décision de transfert des patients en fonction de l'historique de succès des transferts et des profils des organisations concernées.
Support à l'interaction (utilise surtout la reconnaissance)	Interprète et crée du contenu pour alimenter les interactions entre les machines et les humains (p. ex., voix, texte, images).	Patients : Un robot compagnon incluant un agent conversationnel pour briser l'isolement de personnes âgées en hébergement. Cliniciens : Un agent conversationnel répond au clinicien au sujet des caractéristiques de ses patients et des meilleures pratiques cliniques. Gestionnaires : Un moteur de recherche d'entreprise basé sur l'IA permettant d'accéder plus facilement et rapidement au contenu pertinent.
Optimisation centrée sur l'objectif (optimisation mathématique et apprentissage par renforcement)	Cherche la solution optimale selon des critères objectifs, ce qui peut être par apprentissage automatique de type essai-erreur (par renforcement).	Patients : Offrir divers types d'alertes et d'interventions éducatives ajustées en fonction de l'effet observé du traitement. Cliniciens : Aide à la décision clinique qui sélectionne le meilleur moment pour débuter un traitement (p. ex., VIH) pour maximiser les QALYs (<i>quality-adjusted life years</i>). Gestionnaires : Optimisation de la livraison du matériel médical.

Note : Il ne s'agit pas d'une liste exhaustive. D'autres tâches sont possibles. Bien que les tâches soient tirées du cadre de classification de l'OCDE (2022, p. 50), les exemples ont été préparés par l'équipe de travail de la présente fiche de sorte à être contextualisés au secteur de la santé, pour les patients, les cliniciens et les gestionnaires.

Niveau d'autonomie de l'action : selon le rapport de l'OCDE, quatre niveaux permettent de catégoriser l'autonomie d'un SIA, c'est-à-dire dans quelle mesure il peut agir sans intervention humaine; **Aucune** (le SIA ne peut pas agir sur ses propres recommandations et l'humain prend toutes les décisions, que ce soit de considérer ou d'ignorer les résultats du SIA), **faible** (le SIA pourra agir en fonction de ses recommandations, seulement si l'humain approuve l'action), **modérée** (le SIA agira en fonction de ses recommandations, sauf si l'humain refuse), ou élevée (le SIA agit sans implication humaine).

5 RECOMMANDATIONS ENTOURANT L'ADOPTION D'UNE DÉFINITION UTILE DE L'IA EN SANTÉ ET AUTRES PISTES DE RÉFLEXION

À la lumière de ce qui précède, il semble qu'une définition formelle de l'IA n'est peut-être pas aussi utile qu'une description détaillée de ses composantes et de ses caractéristiques. L'adoption d'une définition englobante qui est à la fois claire, flexible et utile est un objectif difficilement atteignable sans passer par un système de classification détaillé. Si une définition formelle de l'IA est tout de même jugée pertinente, la définition suivante, préparée par le groupe d'experts de haut niveau sur l'IA (GEHN IA), puis adoptée par AI Watch (Samoili et al., 2021), est recommandée.

« Les systèmes d'intelligence artificielle (IA) sont des systèmes logiciels (et éventuellement matériels) conçus par des êtres humains et qui, ayant reçu un objectif complexe, agissent dans le monde réel ou numérique en percevant leur environnement par l'acquisition de données, en interprétant les données structurées ou non structurées collectées, en appliquant un raisonnement aux connaissances, ou en traitant les informations, dérivées de ces données et en décidant de la (des) meilleure(s) action(s) à prendre pour atteindre l'objectif donné. Les systèmes d'IA peuvent soit utiliser des règles symboliques ou apprendre un modèle numérique, et peuvent également adapter leur comportement en analysant la mesure dans laquelle l'environnement est affecté par leurs actions préalables. » (GEHN IA, 2018a; 2018b)

Cette définition fut adoptée par AI Watch suite à un travail rigoureux de synthèse des définitions proposées dans des politiques nationales et internationales en matière d'IA à travers le monde, mais aussi dans la littérature scientifique et le monde des affaires. Néanmoins, cette définition pourrait ne pas être adaptée à tous les usages, dans tous les contextes, à tous les instants.

Dans le rapport d'AI Watch, des définitions moins détaillées adaptées à certains usages, ainsi qu'une taxonomie et une liste de mots clés importants associés à l'IA sont également proposées. Cette approche est également celle que nous recommandons, c'est-à-dire de sélectionner, puis d'adapter une définition qui sera utile pour servir les objectifs de chaque organisation. Il serait également idéal que d'encourager l'adoption de différentes définitions pour différents contextes et domaines d'utilisation (p. ex., domaines clinique, logistique, éthique, réglementaire, légal), tout en restant cohérent avec la définition globale retenue. Des outils complémentaires, tels qu'une taxonomie et une liste de mots clés regroupés par catégorie, seraient également pertinents afin d'aider les différents acteurs à déterminer ce qui est considéré comme étant un SIA.

Autres pistes de réflexion

Tel que mentionné dans la mise en contexte, il demeure difficile de définir l'intelligence artificielle, et la multitude de définitions rapportée dans le rapport d'AI Watch en témoigne. Cette fiche offre des pistes permettant de comprendre en quoi la complexité et la diversité de l'IA peuvent en partie expliquer l'absence d'une définition universellement acceptée ou universellement appropriée, ainsi que la pertinence de poursuivre la quête et le développement de définitions non pas universelles, mais utiles.

L'apprentissage automatique rendu possible par les modèles d'IA de type statistique a permis de développer de nouveaux outils et engendrer de nouveaux cas d'usage dans les différents secteurs industriels dont la santé. C'est également principalement cette branche, axée sur les données massives, qui a soulevé de nouveaux enjeux éthiques et légaux. Pour certains milieux, dont celui de la santé, ces enjeux ont une importance capitale.

Le caractère boîte noire de l'IA statistique, c'est-à-dire son caractère peu explicable et interprétable, la prise de décision automatique et le risque d'exacerbation des inégalités sociales en santé associé aux biais algorithmiques ne sont que quelques-unes des thématiques qui soulèvent des préoccupations légales et éthiques importantes. Il est donc fortement conseillé d'entreprendre une réflexion approfondie de ce que représente la valeur ajoutée des SIA en santé (fiche 2) pour en guider un usage responsable (fiches 3 et 4).

6 EN SAVOIR PLUS

La présente section présente sous forme de tableau un groupement thématique de l'ensemble des références citées dans le texte, complétées d'autres références que le lecteur intéressé à en apprendre davantage pourra trouver pertinentes.

Tableau 3. Groupement thématique des références et lectures complémentaires

Thématiques	Références
Histoire de l'IA	<ul style="list-style-type: none"> • Haenlein & Kaplan (2019) • Kaul et al. (2020)
Introduction à l'IA et à l'apprentissage automatique	<ul style="list-style-type: none"> • Beam & Kohane (2018) • Doshi-Velez & Perlis (2019) • Kueper (2021) • Liu et al. (2019) • Rajkomar et al. (2019) • Sutton et al. (2020)
Définitions de l'IA	<ul style="list-style-type: none"> • Groupe d'experts de haut niveau sur l'intelligence artificielle (2018) • High-level expert group on artificial intelligence (2018) • Samoili et al. (2021)
Références techniques sur l'IA	<ul style="list-style-type: none"> • Hastie et al. (2009) • Russell & Norvig (2021) • Sarker & Newton (2007)
Cadres de classification de l'IA	<ul style="list-style-type: none"> • Aiken (2021) • Kaplan & Haenlein (2019; 2022) • OCDE (2022; le document complet et le résumé)
Niveaux de Maturité de la Technologie (NMT)	<ul style="list-style-type: none"> • Cheng et al. (2021) • Innovation Canada (2018) • International Organization for Standardization (2013) • Mankins (1995, 2009) • MEDTEQ+ (2021) • OCDE (2022)
Préoccupations entourant l'IA en santé	<ul style="list-style-type: none"> • Garcia et al. (2020) • Groupe de travail sur l'IA au service de la santé (2020) • Reznick et al. (2020)
Guides d'accompagnement	<ul style="list-style-type: none"> • Li et al. (2020) • Mesko & Gorog (2020) • Tang et al. (2018)
Choisir une définition utile pour faciliter l'implantation	<ul style="list-style-type: none"> • Alsheiabni et al. (2019) • Sadiq et al. (2021) • Schuster et al. (2021)

RÉFÉRENCES

Addo, P., Clark, J., Erdelyi, O., Gomez Gutiérrez, E., Grobelnik, M., Lane, M., Murdick, D., Perset, K., Plonk, A., & Sorgner, A. (2021). *The OECD AI Systems Classification Framework: progress, challenges and way forward*. <https://wp.oecd.ai/app/uploads/2021/02/OECD-AI-Systems-Classification-Framework.pdf>

Aiken, C. (2021). *Classifying AI Systems; CSET Data Brief*. Center for Security and Emerging Technology (CSET). <https://cset.georgetown.edu/wp-content/uploads/Classifying-AI-Systems.pdf>

Alsheiabni, S., Cheung, Y., & Messom, C. (2019). Towards an artificial intelligence maturity model: from science fiction to business facts.

Beam, A. L., & Kohane, I. S. (2018). Big Data and Machine Learning in Health Care. *JAMA*, 319(13), 1317-1318. <https://doi.org/10.1001/jama.2017.18391>

Cheng, W., Xing, C., & Song, G. (2021). A contrastive study of standard specification for technology readiness assessment. 32nd Congress of the International Council of the Aeronautical Sciences, Pudong Shangri-La, Shanghai, China.

Doshi-Velez, F., & Perlis, R. H. (2019). Evaluating Machine Learning Articles. *JAMA*, 322(18), 1777-1779. <https://doi.org/10.1001/jama.2019.17304>

Garcia, J. F., Spatharou, A., Hieronimus, S., Beck, J.-P., & Jenkins, J. (2020). *Transforming healthcare with AI: The impact on the workforce and organizations*. McKinsey & Company, EIT Health. <https://eithealth.eu/wp-content/uploads/2020/03/EIT-Health-and-McKinsey-Transforming-Healthcare-with-AI.pdf>

Groupe d'experts de haut niveau sur l'intelligence artificielle (GEHN IA). (2018). *Lignes directrices en matière d'éthique pour une IA digne de confiance*. Commission européenne. <https://www.justice-ia.com/files/sites/181/2019/10/EthicsguidelinesfortrustworthyAI-FRpdf.pdf>

Groupe de travail sur l'IA au service de la santé (AI4H). (2020). *Bâtir un système de santé apprenant pour les Canadiens; Rapport du groupe de travail sur l'intelligence artificielle au service de la santé*. CIFAR. <https://cifar.ca/wp-content/uploads/2020/11/AI4Health-report-ENG-10-F.pdf>

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 5-14. <https://doi.org/10.1177/0008125619864925>

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.

High-level expert group on artificial intelligence (AI HLEG). (2018). *A definition of AI: Main capabilities and scientific disciplines*. European Commission. https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf

Innovation Canada. (2018). Technology readiness levels. <https://www.ic.gc.ca/eic/site/080.nsf/eng/00002.html>

- International Organization for Standardization. (2013). *Systèmes spatiaux – Definition des Niveaux de Maturité de la Technologie (NMT) et de leurs critères d'évaluation (ISO 16290:2013)*. <https://www.iso.org/fr/standard/56064.html>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, 92(4), 807-812. <https://doi.org/10.1016/j.gie.2020.06.040>
- Kueper, J. K. (2021). Abécédaire de l'intelligence artificielle dans les soins primaires. *Le Médecin de famille canadien*, 67(12), e317-e322. <https://doi.org/10.46747/cfp.6712e317> (Abecedaire de l'intelligence artificielle dans les soins primaires.)
- Li, R. C., Asch, S. M., & Shah, N. H. (2020). Developing a delivery science for artificial intelligence in healthcare. *NPJ Digit Med*, 3, 107. <https://doi.org/10.1038/s41746-020-00318-y>
- Liu, Y., Chen, P. C., Krause, J., & Peng, L. (2019). How to Read Articles That Use Machine Learning: Users' Guides to the Medical Literature. *JAMA*, 322(18), 1806-1816. <https://doi.org/10.1001/jama.2019.16489>
- Mankins, J. C. (1995). *Technology Readiness Level, A White Paper*. NASA.
- Mankins, J. C. (2009). Technology readiness assessments: A retrospective. *Acta Astronautica*, 65(9-10), 1216-1223. <https://doi.org/10.1016/j.actaastro.2009.03.058>
- MEDTEQ+. (2021). Échelle TRL (technology readiness level). <https://www.medteq.ca/ressources/>
- Mesko, B., & Gorog, M. (2020). A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit Med*, 3, 126. <https://doi.org/10.1038/s41746-020-00333-z>
- Murdick, D. (2020). *AI System Classification for Policymakers; ONE AI WG on AI System Classification*. <https://www.oecd.org/parliamentarians/meetings/ai-meeting-december-2020/>
- OECD. (2022). *An overview of the OECD Framework for Classifying AI systems*. OECD.AI Policy Observatory. <https://wp.oecd.ai/app/uploads/2022/02/Classification-2-pager-1.pdf>
- Organisation for Economic Co-operation and Development (OECD). (2022). *OECD framework for the classification of AI systems – Public consultation on preliminary findings*. <https://oecd.ai/en/classification>
- Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14), 1347-1358. <https://doi.org/10.1056/NEJMr1814259>
- Reznick, R. K., Harris, K., Horsley, T., & Hassani, M. S. (2020). Task force report on artificial intelligence and emerging digital technologies. *R Coll Physicians Surg Canada*, 1-52.
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th Edition, Global Edition ed.). Pearson Education Limited.
- Sadiq, R. B., Safie, N., Abd Rahman, A. H., & Goudarzi, S. (2021). Artificial intelligence maturity model: a systematic literature review. *PeerJ Comput Sci*, 7, e661. <https://doi.org/10.7717/peerj-cs.661>

Samoili, S., López Cobo, M., Delipetrev, B., Martínez-Plumed, F., Gómez, E., & De Prato, G. (2021). *AI watch, defining Artificial Intelligence 2.0; Towards an operational definition and taxonomy for the AI landscape* (Vol. EUR 30873 EN). Publications Office of the European Union. <https://doi.org/10.2760/019901>

Sarker, R. A., & Newton, C. S. (2007). *Optimization modelling: a practical approach*. CRC press.

Schuster, T., Waidelich, L., & Volz, R. (2021). Maturity Models for the Assessment of Artificial Intelligence in Small and Medium-Sized Enterprises. PLAIS EuroSymposium on Digital Transformation,

Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med*, 3, 17. <https://doi.org/10.1038/s41746-020-0221-y>

Tang, A., Tam, R., Cadrin-Chenevert, A., Guest, W., Chong, J., Barfett, J., Chepelev, L., Cairns, R., Mitchell, J. R., Cicero, M. D., Poudrette, M. G., Jaremko, J. L., Reinhold, C., Gallix, B., Gray, B., Geis, R., & Canadian Association of Radiologists Artificial Intelligence Working, G. (2018). Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology. *Canadian Association of Radiologists Journal*, 69(2), 120-135. <https://doi.org/10.1016/j.carj.2018.02.002>



**OBSERVATOIRE INTERNATIONAL
SUR LES IMPACTS SOCIÉTAUX
DE L'IA ET DU NUMÉRIQUE**

Pavillon Charles-De Koninck, local 2489
1030, avenue des Sciences-Humaines
Université Laval
Québec (Québec) G1V 0A6

observatoireia@ulaval.ca
observatoire-ia.ulaval.ca