Introduction à l'IA en santé publique

Série de conférences de l'ASPC
2026-09-23

Professeure Laura C. Rosella
Chaire de recherche du Canada en transformation et analytique de la santé des populations (niveau 1)
Dalla Lana School of Public Health, University of Toronto
Responsable pédagogique, *Temerty Centre for Artificial Intelligence Research and Education in Medicine (T-CAIREM)*
Candidate principale désignée pour la plateforme de formation à la recherche de *AI for Public Health (AI4PH)*
(https://ai4ph-hrtp.ca/)

AI4PH AI FOR PUBLIC HEALTH

UNIVERSITY OF TORONTO
DALLA LANA SCHOOL OF PUBLIC HEALTH

Aujourd’hui, j’aborderai les points suivants :

Notions de base

Concepts et limites

Utilisations : comment utilise-t-on l’IA?

Tâches et valeur ajoutée

Répercussions

Risques pour la santé, l’équité et l’environnement

Mise en œuvre responsable

Exigences pratiques

Notions de base · Utilisations · Répercussions · Mise en œuvre responsable

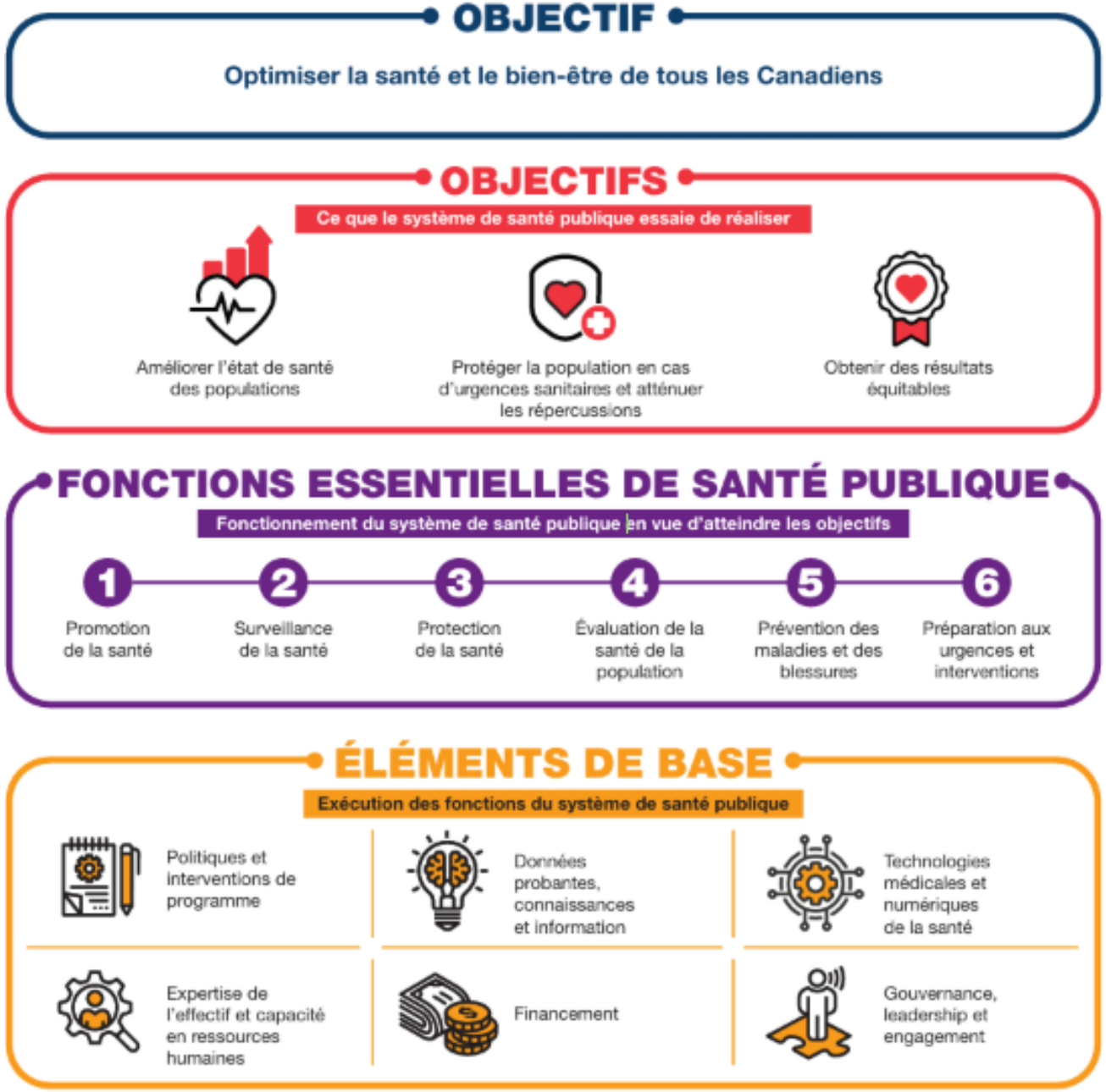
Comment savoir si les progrès de l'IA ont des répercussions positives sur la santé publique?

⇒Il y a moins de gens qui développent des maladies ou qui souffrent de problèmes de santé évitables

⇒Nous faisons la promotion de la santé dans nos environnements et nos communautés

⇒Les iniquités en santé diminuent

Unclassified / Non classifié

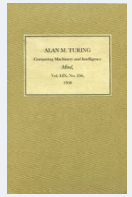


Petit aperçu historique

Partie 1 · 1950 à 1999 : des grandes idées aux grandes victoires

Années 1950

Naissance de l'IA



1950 Turing : « Les machines peuvent-elles penser? »
1956 Dartmouth donne le nom au domaine
1957 Perceptron

Années 1960

Optimisme précoce



Raisonnement heuristique et symbolique
1966 ELIZA, le premier robot conversationnel
1966 Shakey, un robot capable de planifier

Années 1970 et 1980

Hivers de l'IA et systèmes experts

❄

1973 Premier hiver de l'IA
Les systèmes experts **MYCIN** et **XCON**
1987 Deuxième hiver de l'IA

Années 1990

Jeux et statistiques



1992 TD-Gammon : premiers succès en AR
1997 Deep Blue vainc Kasparov
Vision statistique et TALN

Petit aperçu historique - 2

Partie 2 · 2000 à aujourd'hui : données, apprentissage profond et IA générative

Années 2000

Données et calcul



2005 Défi des véhicules sans pilote DARPA
2006 Prix Netflix
2009 Ensemble de données ImageNet

Années 2010

Apprentissage profond



2012 AlexNet
2014 Réseaux antagonistes génératifs, séquence à séquence
2015 ResNet, Réseaux de neurones DQN
2016 AlphaGo
2017 Transformeurs
2018 BERT

Années 2020

IA générative



2020 AlphaFold 2
2022 ChatGPT, Stable Diffusion
2023 LLaMA à poids ouvert
2024 Prix Nobel pour l'IA
Maintenant Agents multimodaux

*English only content / Contenu en anglais seulement

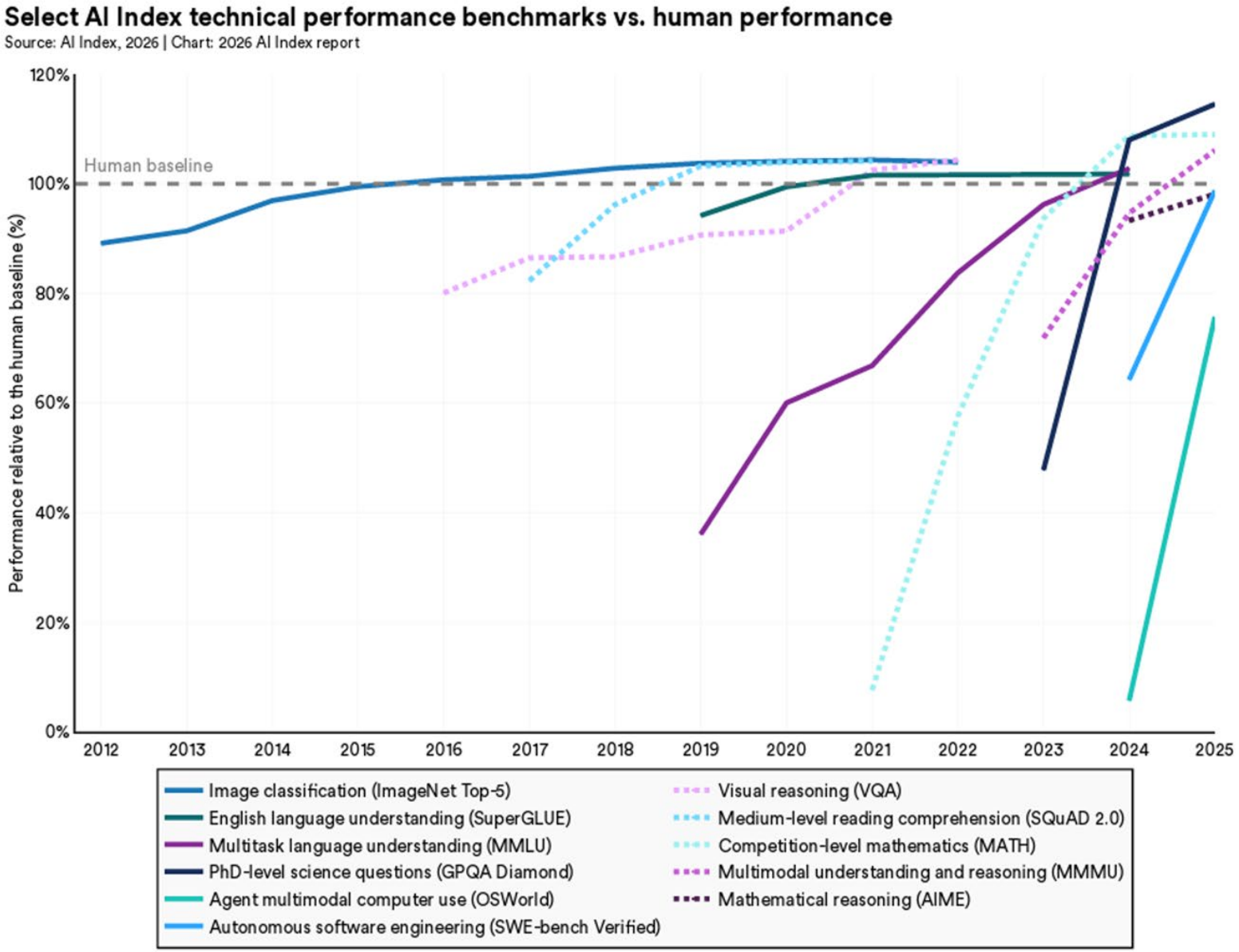
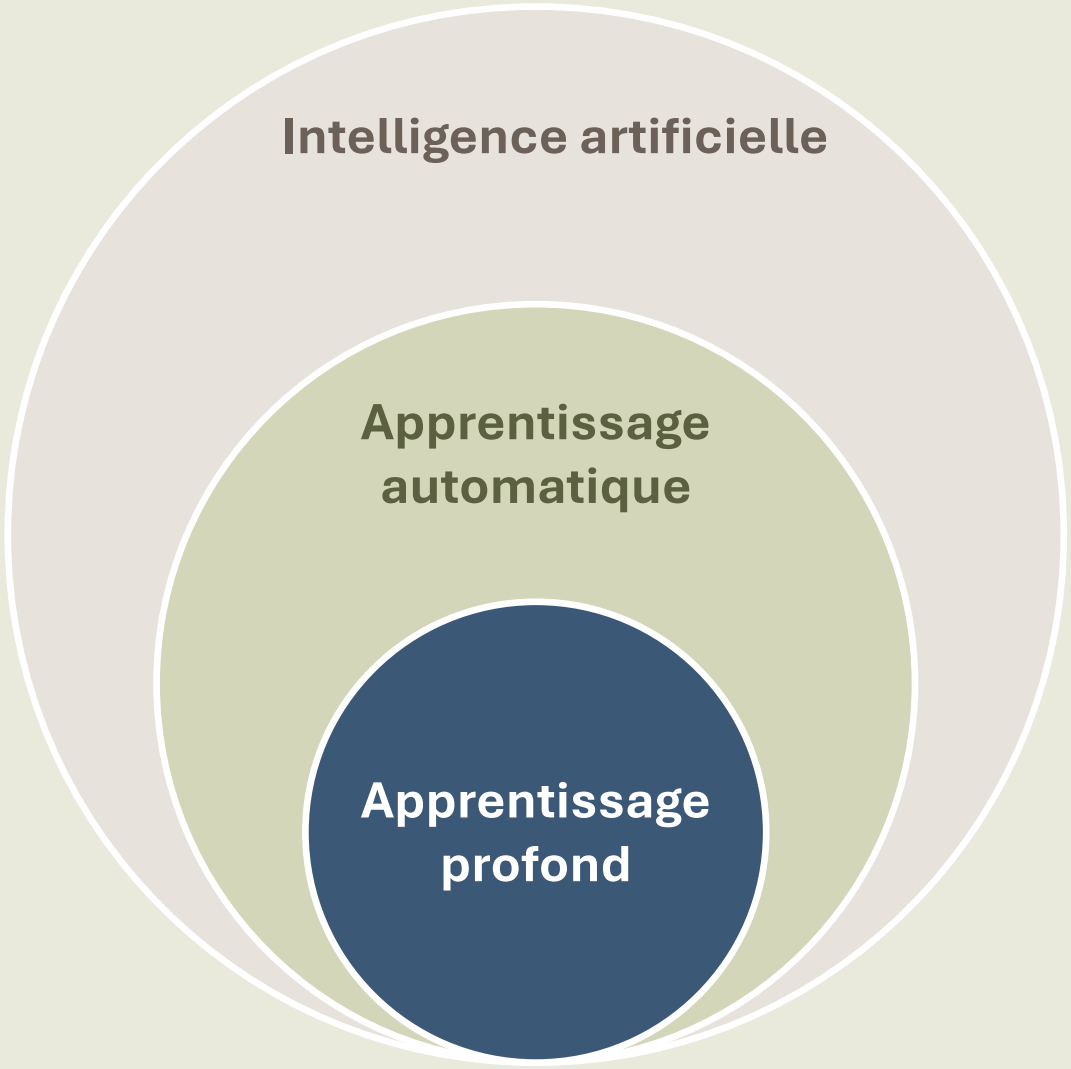


Figure 2.1.1[1]

Rapport de 2026 sur l'indice de l'IA

Stanford University Human-Centered Artificial Intelligence

IA, apprentissage automatique et apprentissage profond



Intelligence artificielle
Des machines effectuent des tâches nécessitant une intelligence semblable à celle d'un humain

Apprentissage automatique
On parle ici de l'apprentissage de modèles à partir de données

Apprentissage profond
On parle ici de réseaux neuronaux à plusieurs couches pour les images, le texte, le son et d'autres données complexes

IA générative

Type d'IA qui apprend des modèles à partir de données d'entraînement pour générer du nouveau contenu en réponse à des instructions.



Fonctionnement

- Pendant l'entraînement, les modèles apprennent des schémas récurrents dans le langage, les images ou d'autres données
- Les grands modèles de langage génèrent du texte par petits morceaux, en utilisant la requête, le contexte et le texte précédent pour prédire la suite
- Les instructions et les documents fournis déterminent la réponse, notamment son contenu, son ton et son niveau de détail

Possibilités en matière de santé publique

- Transformer un rapport technique en ébauche de note d'information ou en explication destinée au public
- Adapter des directives existantes à différents publics, langues ou niveaux de lecture
- Explorer des thèmes dans les réponses de consultations ou générer du code pour faciliter l'analyse
- Vérifier les résultats!

Formation, tests et utilisation concrète

Stade	Question
Formation	Le modèle peut-il apprendre des schémas utiles à partir de ces données?
Validation et tests	Les schémas restent-ils valables dans des données qui n'ont pas servi à l'ajustement du modèle?
Évaluation en conditions réelles	L'utilisation du modèle permet-elle d'améliorer les décisions et les résultats?
Suivi	Reste-t-il utile quand le contexte change?

Apprentissage automatique en santé publique

Cinq approches, de la recherche de schémas à la génération de texte

APPROCHE	UTILITÉ	EXEMPLE DE SANTÉ PUBLIQUE
Regroupement	Identifier les regroupements naturels dans les données	Identification des zones névralgiques et des signes d'éclosion
Arbres décisionnels et ensembles	Prédire un résultat à partir de nombreux facteurs	Anticipation des risques, recherche de cas, prévisions
Apprentissage profond	Identifier des schémas dans les images, signaux et données complexes	Imagerie, capteurs et appareils portables
Traitement automatique du langage naturel	Extraire le sens d'un texte	Identification des cas, symptômes et opinion publique
Grands modèles de langage multimodaux	Générer et combiner du texte, des images et des données	Résumé de rapports; extraction d'informations

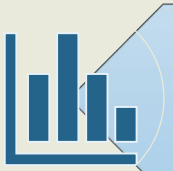
Sources : OMS, 2021; OMS, 2024; Sidey-Gibbons & Sidey-Gibbons, 2019; LeCun et al., 2015; Moor et al., 2023.

*English only content / Contenu en anglais seulement

Tendances technologiques émergentes

- Modèles fondateurs : modèles d'IA polyvalents qui peuvent être adaptés à plusieurs tâches différentes
- IA en temps réel : systèmes qui détectent, se mettent à jour et adaptent leurs réponses en continu
- IA multimodale : systèmes qui combinent des sources de données issues de plusieurs modalités
- IA agentique : systèmes qui planifient et exécutent plusieurs étapes à l'aide d'outils
- Intelligence artificielle générale (IAG) : niveau de performance égal ou supérieur à celui de l'humain pour la plupart des tâches cognitives

Détails sur les applications de l'IA



Volet descriptif



Volet prédictif



Inférence causale



Volet prescriptif

Hernan et al. CHANCE 2019

Détails sur les applications de l’IA – 2



Que se passe-t-il?



Que va-t-il se passer?



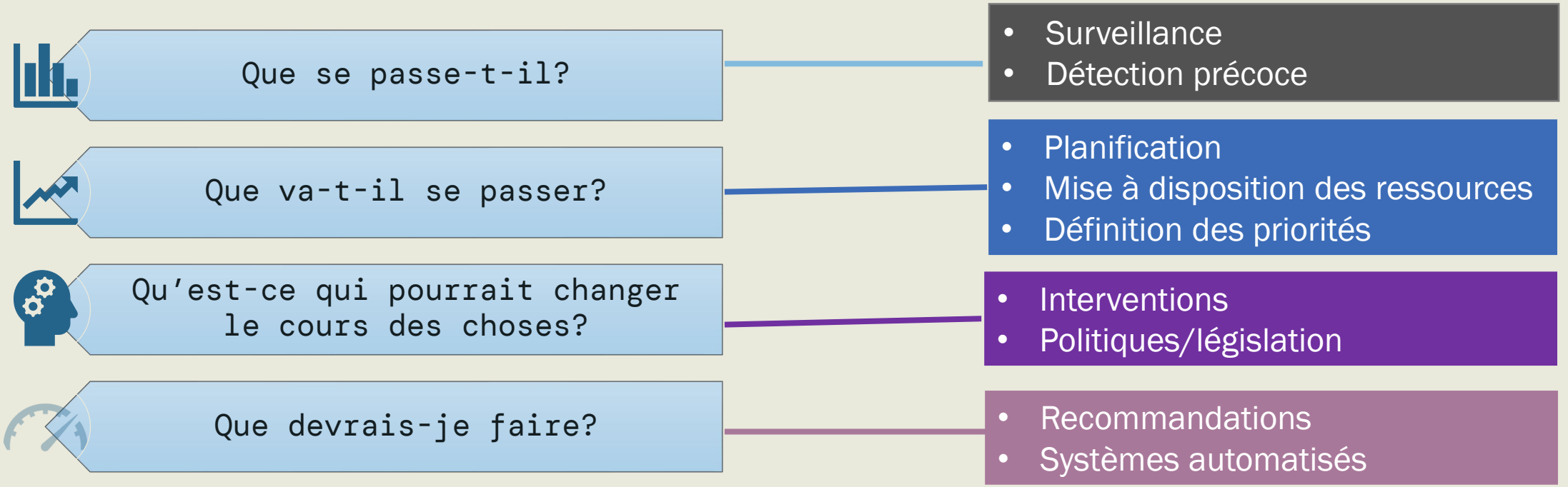
Qu’est-ce qui pourrait changer le cours des choses?



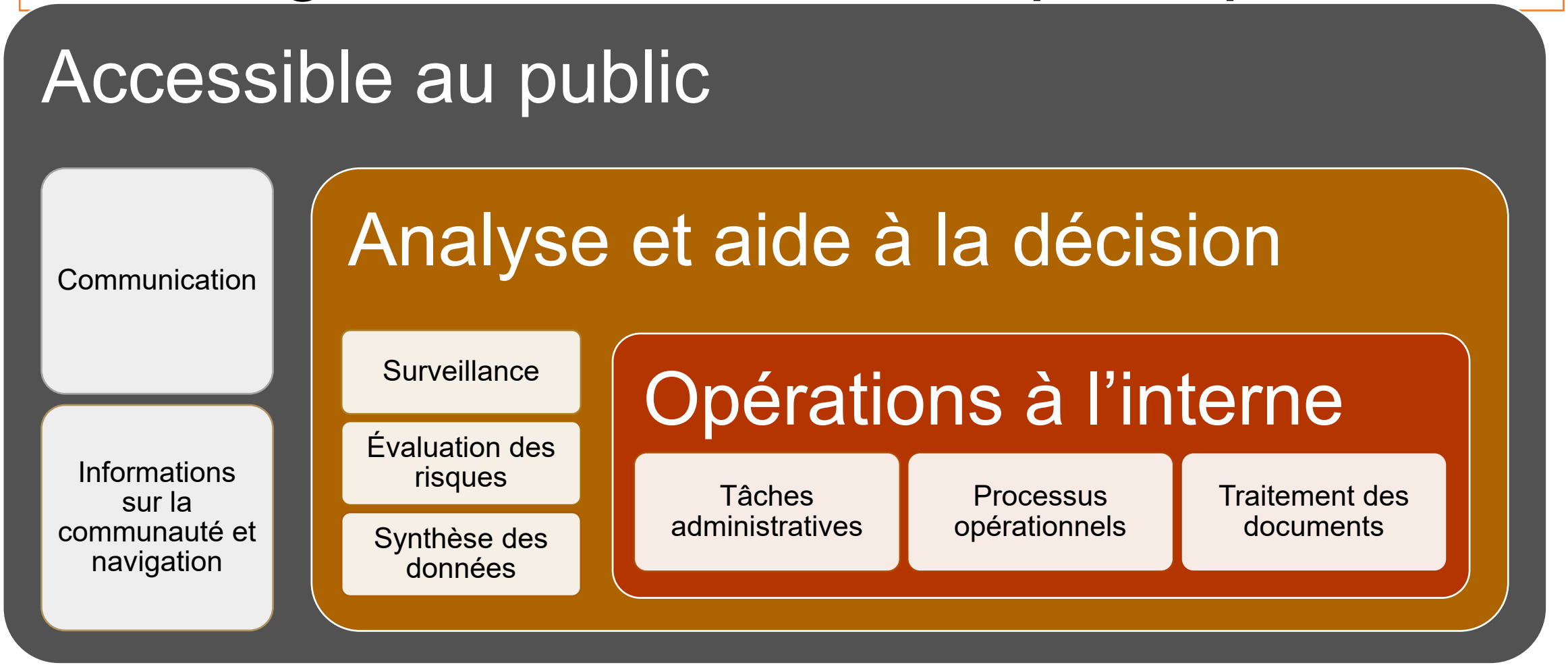
Que devrais-je faire?

Hernan et al. CHANCE 2019

Détails sur les applications de l’IA en santé publique



Quelle est la place de l'IA au sein des organisations de santé publique?



Article sur la classification de textes pour les enquêtes sur les éclosions de toxi-infections d'origine alimentaire à l'aide de l'apprentissage profond

- **Tâche :** transformer les historiques alimentaires sous forme de texte libre en catégories alimentaires standardisées pour l'analyse des éclosions
- **Contribution de l'IA :** un modèle de langage attribue des catégories alimentaires à des descriptions non structurées, ce qui aide les responsables d'enquêtes à analyser les informations d'exposition à grande échelle
  - Grâce au calcul par le biais d'un réseau de systèmes informatiques, le *CDC* a entraîné un **réseau neuronal profond**, un modèle de langage servant à prédire les catégories alimentaires à partir d'observations sous forme de texte libre

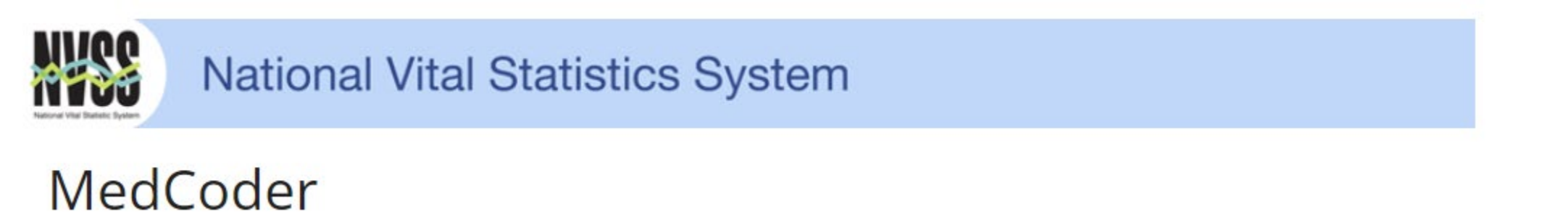
Volet descriptif

Surveillance

Khan, MA. et al. *Text Classification for Foodborne Illness Outbreak Investigations Using Deep Learning* EIS Conference 2023

*English only content / Contenu en anglais seulement

Exemple du *CDC* : MedCoder



- **Tâche** : convertir les causes de décès indiquées sur les actes de décès en code CIM-10 normalisés
- **Contribution de l'IA** : traitement du langage et apprentissage automatique, combinés à des règles de codage
  - *D'après le CDC, MedCoder peut coder automatiquement près de 90 % des dossiers, contre moins de 75 % pour le système précédent*

Volet descriptif

Surveillance

CDC. *A Timelier Understanding of Death*. 2023
https://stacks.cdc.gov/view/cdc/133554

Research

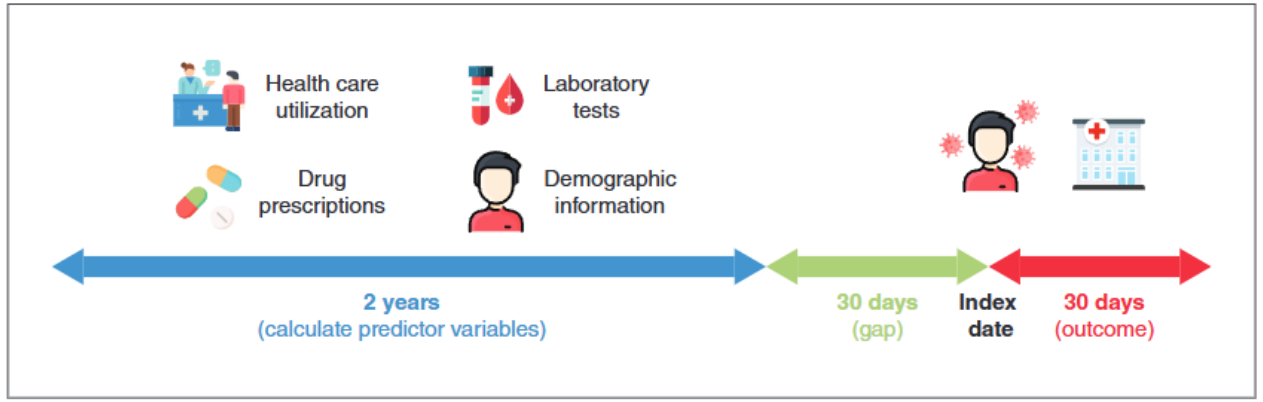
Risk stratification for COVID-19 hospitalization: a multivariable model based on gradient-boosting decision trees

Jahir M. Gutierrez PhD, Maksims Volkovs PhD, Tomi Poutanen MSc, Tristan Watson MPH, Laura C. Rosella PhD

Volet prédictif

*English only content / Contenu en anglais seulement

- **Tâche** : estimer le risque d'hospitalisation dans les 30 jours suivant un test positif au SRAS-CoV-2
- **Contribution de l'IA** : un modèle d'amplification du gradient combine des données de santé recueillies régulièrement pour identifier les personnes à risque élevé plus efficacement que les règles fondées sur l'âge et les comorbidités évaluées



Définition des priorités

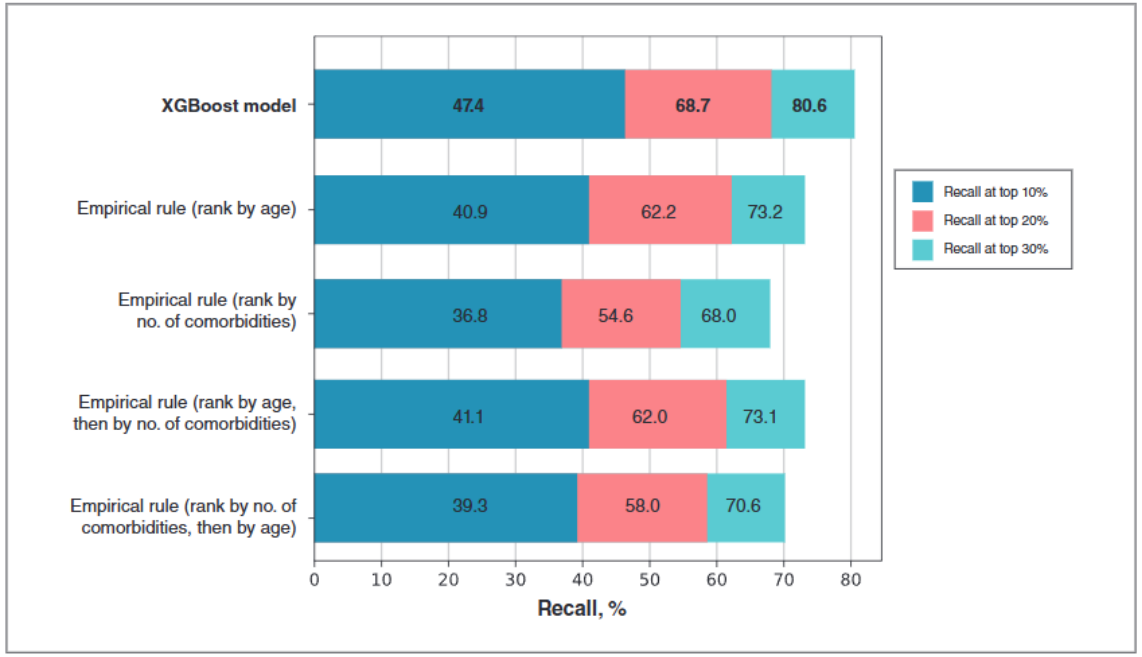


Figure 4: Comparison of recall at top percentiles. The final Extreme Gradient Boosting (XGBoost) model recall (percentage of true hospitalizations recovered in validation data set) was compared against 4 empirical rules.

Gutierrez et al. CMAJ Open Oct 2021, 9 (4)

*English only content / Contenu en anglais seulement

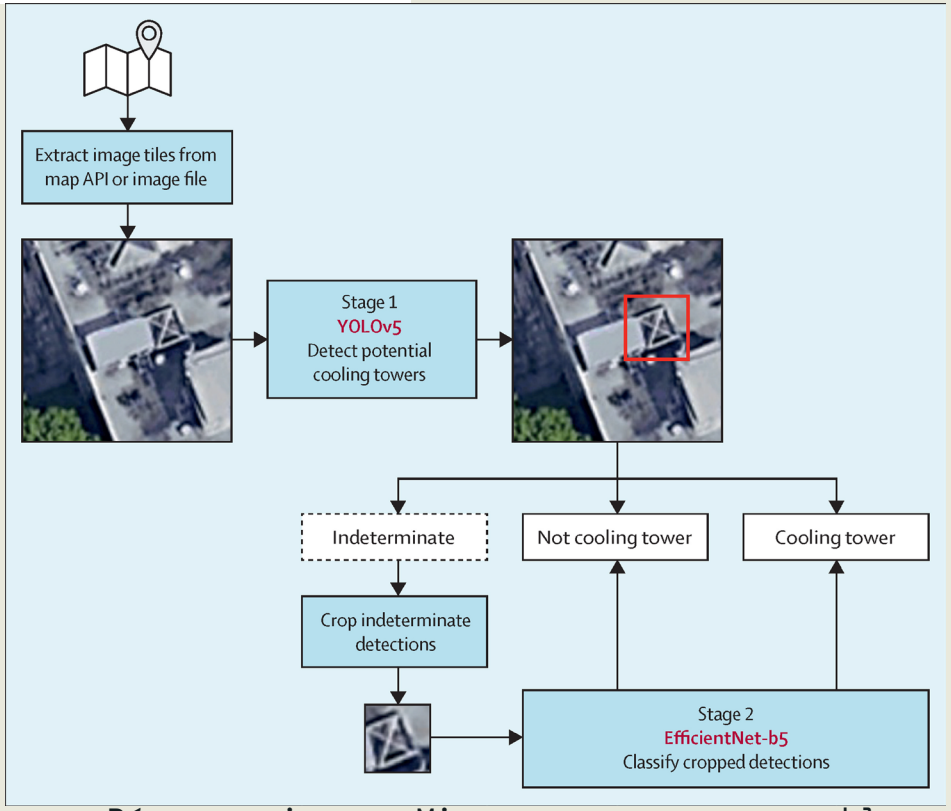
Surveillance

Définition des priorités

Automated cooling tower detection through deep learning for Legionnaires' disease outbreak investigations: a model development and validation study

Karen K Wong, Thaddeus Segura, Gunnar Mein, Jia Lu, Elizabeth J Hannapel, Jasen M Kunz, Troy Ritter, Jessica C Smith, Alberto Todeschini, Fred Nugen, Chris Edens

- **Tâche** : repérer les tours de refroidissement qui pourraient nécessiter une enquête en tant que sources potentielles d'exposition à *Legionella*
- **Contribution de l'IA** : la vision par ordinateur analyse les images aériennes à la recherche de tours de refroidissement potentielles, ce qui permet de cibler la recherche pour les responsables d'enquêtes et de réduire l'examen manuel des images



Volet descriptif

*English only content / Contenu en anglais seulement

Volet descriptif

Planification des communications

PLOS DIGITAL HEALTH

RESEARCH ARTICLE

Use of large language models as a scalable approach to understanding public health discourse

Laura Espinosa*, Marcel Salathé

Digital Epidemiology Lab, School of Life Sciences, School of Computer and Communication Sciences, EPFL, Switzerland

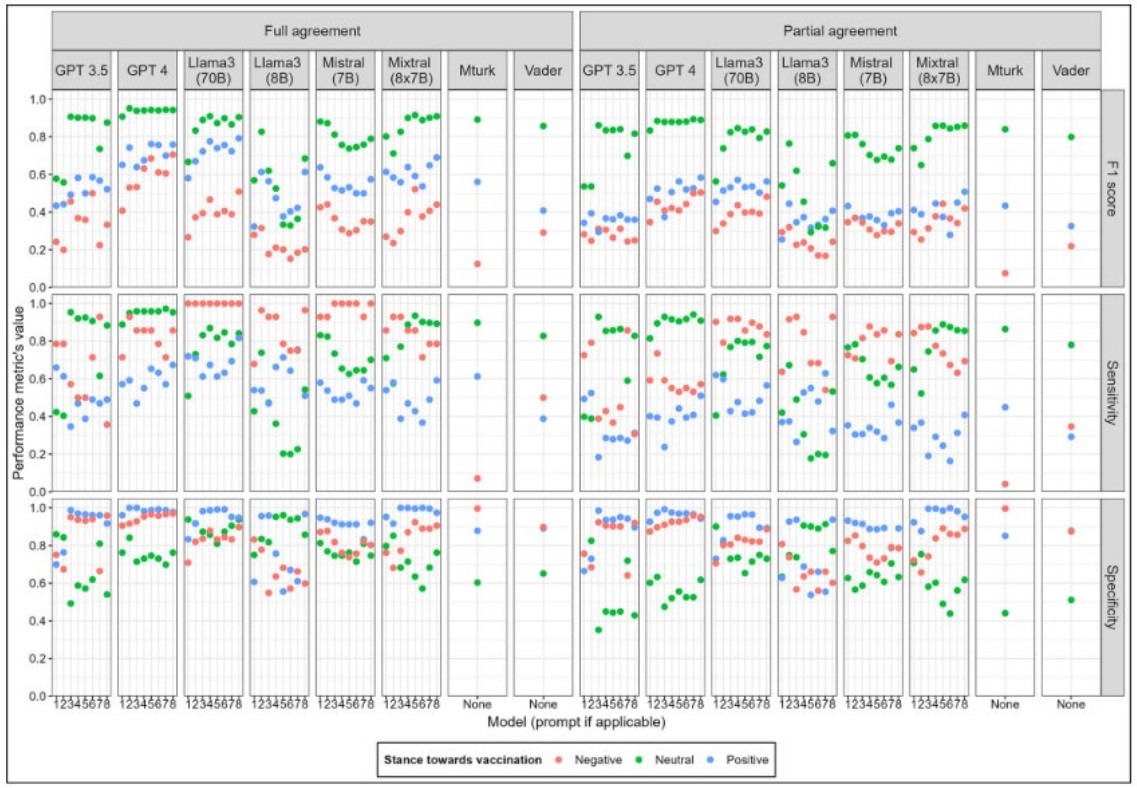


Fig 3. F1 score, sensitivity and specificity per model, prompt (if applicable), stance and experts' agreement.

- **Tâche :** déterminer si les publications sur les réseaux sociaux expriment une position positive, négative ou neutre par rapport à la vaccination
- **Contribution de l'IA :** les modèles de langage préentraînés classent les publications non structurées à l'aide d'instructions et d'exemples, offrant ainsi une solution de remplacement évolutive au marquage manuel de chaque publication

Enseignements tirés de ces exemples

- L’IA permet de structurer les informations, d’identifier les schémas et de faciliter la définition des priorités.
- Chaque résultat doit être clairement lié à une mesure de santé publique.
- L’évaluation doit distinguer les performances du modèle, ses effets sur le travail et ses conséquences sur la santé.
- Les avantages et les inconvénients doivent être évalués à l’échelle de l’ensemble des populations.

Défi : garantir la pertinence des actions

En tant que [décisionnaire en santé publique],
si je savais [information],
je ferais [intervention/action],
pour améliorer [résultat mesurable]

Qualité, représentativité et adéquation des données à l’usage prévu

La façon dont les données sont recueillies, mesurées et utilisées détermine ce que l’IA peut nous révéler

Questions à poser – Nouveauté, représentativité, rapidité, longévité



Nouveauté : L'utilisation de ces données permettra-t-elle d'obtenir de nouvelles informations?



Représentativité : Qui ces données représentent-elles?



Rapidité : L'utilisation de ces données permettra-t-elle d'obtenir des informations plus rapidement?



Longévité : Peut-on conserver ces données pour des applications en cours?

Sources de données émergentes – exemples

Données sur les eaux usées

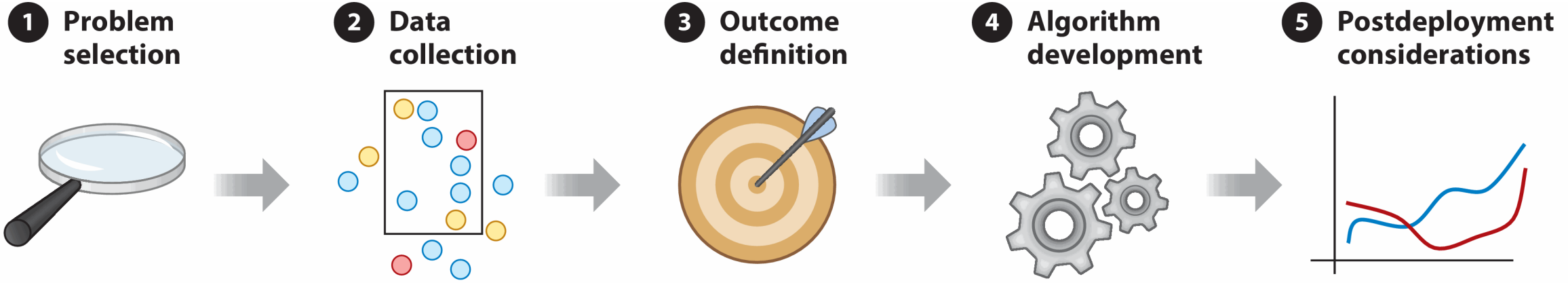
- ✓Rapidité
- ✓Couverture de la population
- ✓Risques faibles sur le plan juridique et de la confidentialité
- ✓Exploitation
- ❖Processus coûteux?

Données sur les réseaux sociaux

- ✓Rapidité
- ✓Rentabilité
- ✓Nouveauté des informations
- ❖Biais de sélection
- ❖Enjeux de confidentialité
- ❖Difficultés quant à la mise en œuvre

Unclassified / Non-classifié

*English only content / Contenu en anglais seulement



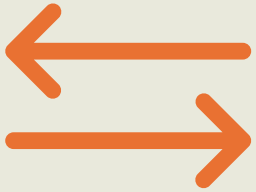
Chen IY, et al. 2021
Annu. Rev. Biomed. Data Sci. 4:123–44

Répercussions de l’IA sur l’environnement

Stade	Ressources et répercussions à prendre en compte
Fabrication du matériel	Extraction des matières premières, fabrication et émissions intrinsèques
Formation et utilisation	Électricité pour les calculs, y compris les inférences répétées
Infrastructure	Refroidissement, consommation d’eau et approvisionnement local en électricité
Fin de vie	Cycles de remplacement, réutilisation et déchets électroniques

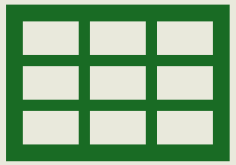
IEA, *Energy and AI*, 2025

Entraînement de nos modèles



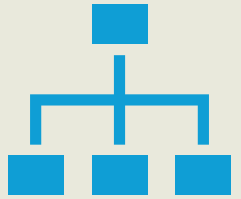
Apprentissage par transfert

Utiliser un modèle préentraîné et l'ajuster pour une nouvelle tâche



Élagage et quantification

Supprimer les paramètres moins importants; réduire la précision numérique



Modèles plus petits et spécifiques à une tâche

Concevoir des modèles spécialisés plutôt que généraux

Flux de travail

Diminuer les formations redondantes

Améliorer l'efficacité des centres de données

Éviter les utilisations inutiles

Gouvernance de l’IA

Politiques	Institutions	Technique
• Réglementation des centres de données • Réglementation sur l’eau • Responsabilité en matière de déchets électroniques	• Normes d’approvisionnement • Gestion des calculs • Affinement des tâches	• Limites de taille des modèles • Rapports sur les émissions de carbone • Apprentissage par transfert par défaut

Unclassified / Non classifié

Conditions pour une mise en œuvre réussie en santé publique

Harmonisation avec les objectifs de santé publique

- Diminution du nombre de personnes qui développent des maladies ou qui souffrent de problèmes de santé évitables
- Promotion de la santé dans nos environnements et nos communautés
- Réduction des iniquités en santé
- Durabilité

Efficacité

- Test et vérification de manière indépendante
- Cohérence et précision dans des conditions réelles
- Adaptation à différentes populations
- Supervision et adaptation aux changements et aux mises à jour du modèle

Sécurité et éthique

- Confidentialité et sécurité des données
- Contrôle humain des décisions
- Conformité juridique et réglementaire
- Lutte contre les préjugés et la discrimination; accès équitable

Six domaines prioritaires pour que les organisations de santé publique mettent l'IA en pratique*

Gouvernance

Infrastructure

Effectifs

Partenariats

Pratiques exemplaires en matière d'IA

Équité et biais

Cela comprend les systèmes sécurisés qui prennent en charge l'intégration, les tests, l'entretien et la surveillance**

Les responsables de l'examen doivent disposer de l'expertise, du temps et de l'autorité nécessaires pour remettre en question les résultats et prendre des mesures face aux problèmes**

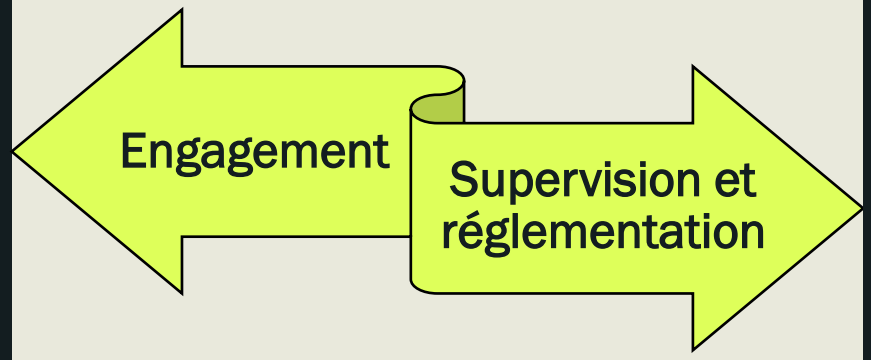
*Fisher S, Rosella LC. BMC Public Health. 2022;22:2146

**Hattab G, et al. AJPH 2025;115(2):123–128.

Construction d’une IA responsable pour la santé

Accent sur les personnes, les praticiens et praticiennes ainsi que les responsables des opérations

- Toutes les étapes : identification du problème, conception, mise en œuvre, évaluation
- Réponse aux véritables défis de santé publique et communautaires
- Appui sur des relations authentiques

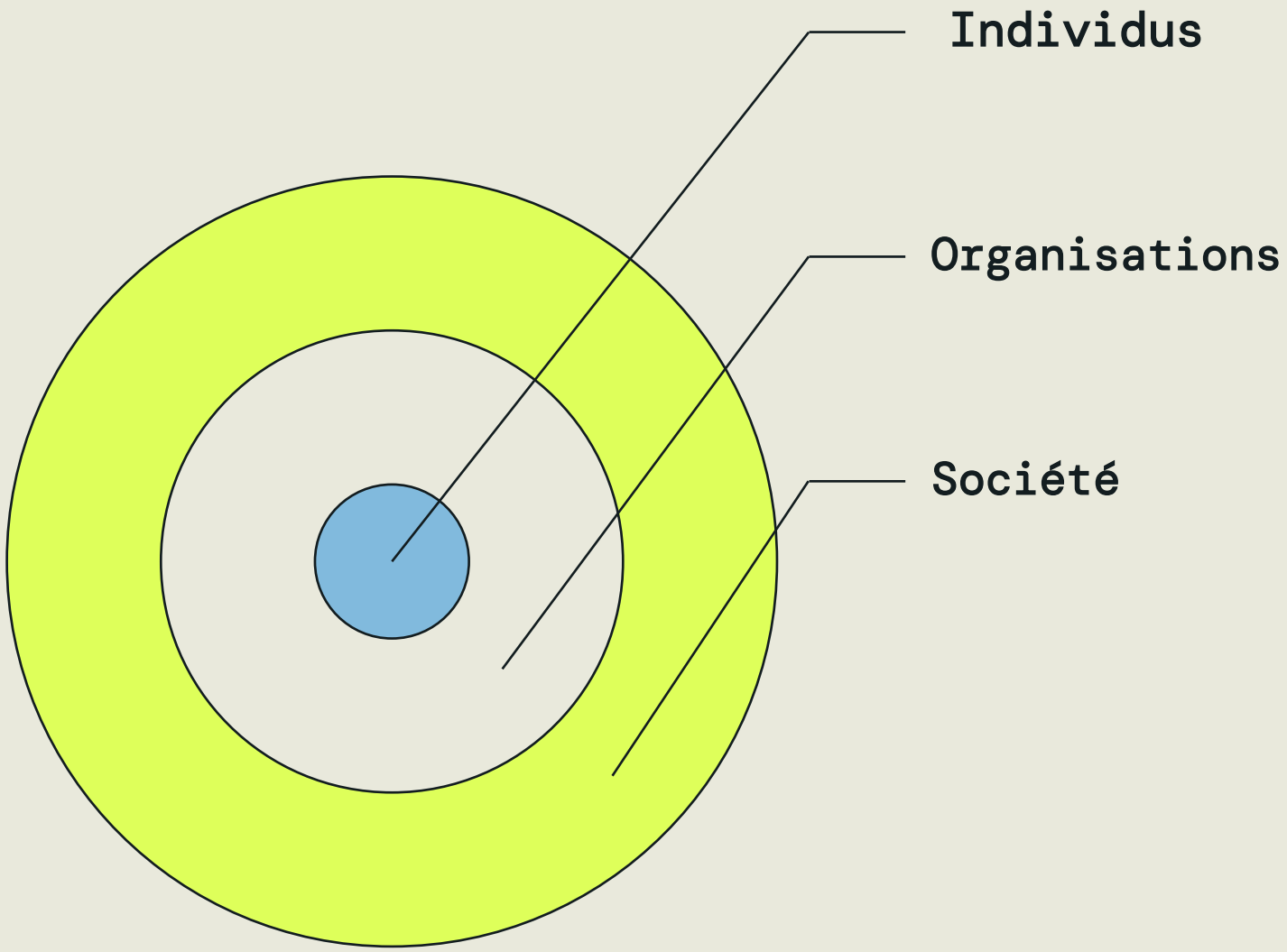


Réglementation : Règles ou lois officielles établies par un gouvernement ou un organisme de réglementation

Directives : lignes directrices ou fortes recommandations sans obligation d’application

Normes : documents officiels fixant des critères et des exigences spécifiques; plus détaillées que des directives

Comment faire?



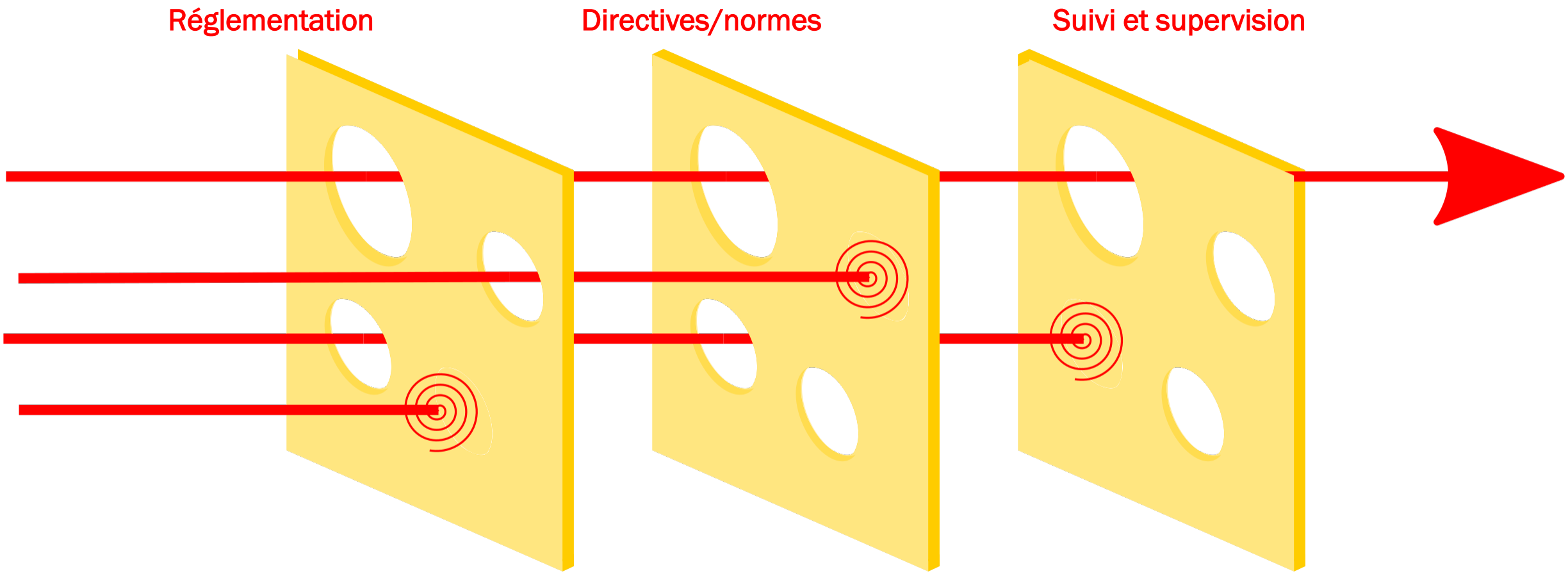


Image Adapted from BenAveling, Wikimedia Commons, CC BY-SA 4.0a

OMS : six principes pour l'IA en santé

1 Protéger l'autonomie de l'être humain

Les êtres humains doivent rester maîtres des décisions relatives à la santé; la vie privée et le consentement sont respectés

2 Promouvoir le bien-être et la sécurité des personnes et l'intérêt public

Les outils répondent aux exigences en matière de sécurité, d'exactitude et d'efficacité, et préviennent les préjudices

3 Garantir la transparence, l'explicabilité et l'intelligibilité

Les technologies sont compréhensibles pour les développeurs, les utilisateurs et le public, et peuvent être remises en question

4 Favoriser la responsabilité et la responsabilisation

La responsabilité face aux résultats est claire, avec des moyens de contester les décisions et de demander réparation

5 Garantir l'inclusion et l'équité

L'IA est conçue de manière à encourager une utilisation et un accès équitables, indépendamment de l'âge, du sexe, du revenu, de la race, des capacités ou d'autres caractéristiques

6 Promouvoir une intelligence artificielle réactive et durable

Les conditions réelles d'utilisation sont évaluées en continu, en tenant compte des répercussions sur l'environnement et le personnel

Organisation mondiale de la Santé. Éthique et gouvernance de l'intelligence artificielle pour la santé : orientation de l'OMS. Genève : OMS; 2021.

Mobilisation pour faire avancer l'intelligence artificielle dans la prédiction et la prévention du diabète

*English only content / Contenu en anglais seulement



www.instituteforbetterhealth.com/ai4dpp-meeting/



https://cifar.ca/fr/evenements/reunion-du-reseau-de-solutions-lia-au-service-de-la-prediction-et-de-la-prevention-du-diabete/

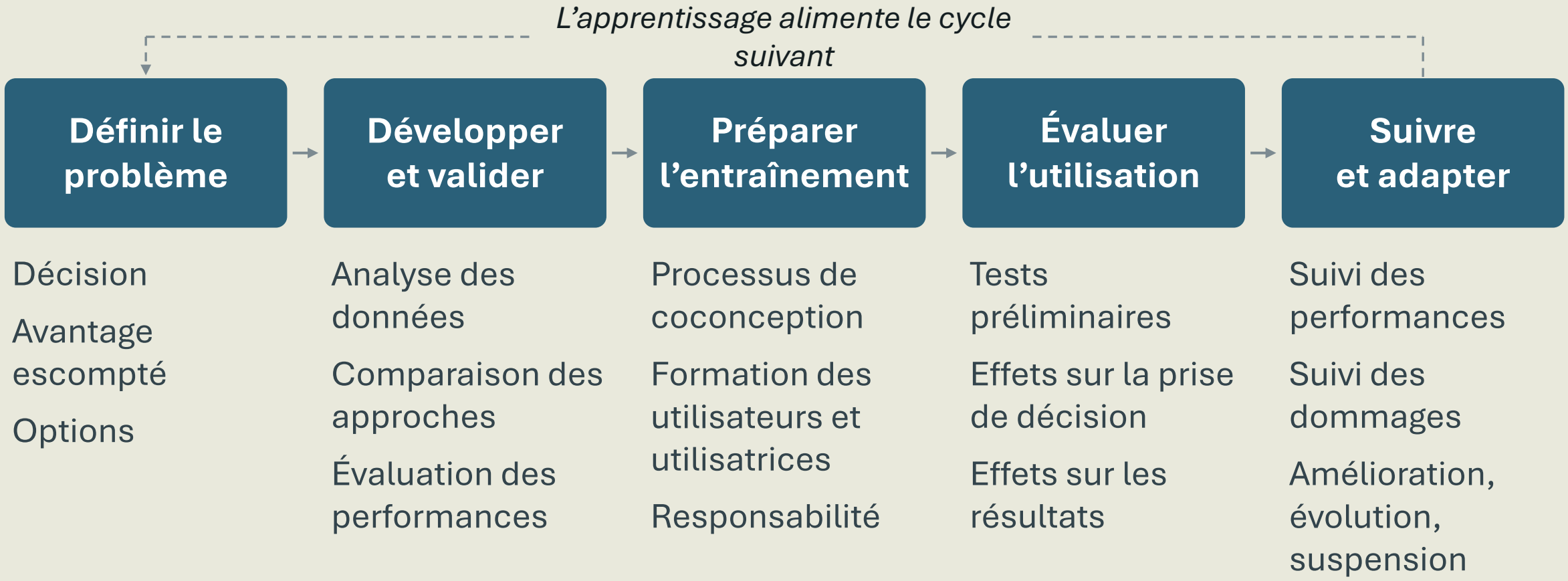
- **Tâche :** mettre au point, avec les praticiens et praticiennes ainsi que les partenaires communautaires, une approche utile et acceptable pour la prédiction et la prévention du diabète
- **Contribution de l'IA :** estimations de risque pour aider à hiérarchiser les efforts de prévention
- **Avantage de la participation :** décider quels résultats sont utiles, comment ils doivent s'articuler avec les services, comment ils sont communiqués et ce qu'implique une utilisation responsable

1. **Se réunir, apprendre et échanger des idées sur l'IA et la santé**
2. **Concevoir ensemble le déploiement de l'IA avec les praticiens, les praticiennes et la communauté**
3. **Contribuer aux résultats de l'IA pour le diabète dans la région de Peel**

Abejirinde IO et al. J Particip Med. 2025 Oct 2;17.

Rosella LC, et al. *DIGITAL HEALTH*. 2025;11

Cycle de vie de la mise en œuvre responsable



Participation communautaire, équité et gouvernance continues

Que devrait-on évaluer?

Dimension	Question clé
Performance	L'outil fonctionne-t-il de manière fiable auprès de la population et dans le contexte visé?
Pratique	Améliore-t-il la prise de décision ou réduit-il la charge de travail, une fois pris en compte la vérification et la correction?
Avantages pour la santé publique	Son utilisation améliore-t-elle la prévention, la rapidité d'intervention ou les résultats en santé?
Équité	Qui en bénéficie, qui en est exclu et qui subit des dommages?
Viabilité	Les avantages justifient-ils les coûts, le besoin de main-d'œuvre et les répercussions sur l'environnement?

Rosella LC. Will AI improve public health? Setting the standards for meaningful impact. *CJPH.* 2025;116:815–818.

Notre rôle pour façonner l’avenir

Notions de base

Comprendre les capacités de l'IA… et ses limites

Utilisations

Commencer par les objectifs et les fonctions de santé publique, puis voir ce que l'IA peut apporter

Répercussions

Faire en sorte que la protection des personnes, l'équité et l'environnement fassent partie du projet

Mise en œuvre responsable

Transformer des outils prometteurs en avantages durables pour la santé publique

Merci!

Laura C. Rosella, PhD
laura.rosella@utoronto.ca
@LauraCRosella
www.linkedin.com/in/laura-rosella-b68b5b91

UNIVERSITY OF TORONTO
DALLA LANA SCHOOL OF PUBLIC HEALTH

Population Health Analytics Lab
www.pophealthanalytics.com
@PopHlthAnalytic

Population Health Analytics Laboratory

AI4PH AI FOR PUBLIC HEALTH

https://ai4ph-hrtp.ca/
@ai4ph
/ai4ph-health-research-training-platform

Unclassified / Non classifié

AI for Public Health (AI4PH)

Mission et approche

MISSION

Constituer un réseau de chercheurs et de chercheuses ainsi que de praticiens et de praticiennes du domaine de la santé publique capables de développer et de mettre en œuvre des innovations en matière d'intelligence artificielle (IA) et d'apprentissage automatique équitables, afin de les appliquer à la recherche, aux politiques et aux pratiques de santé publique au Canada.

APPROCHE

Notre démarche consiste à développer un ensemble d'activités et d'occasions de formation destinées à faciliter la mise en œuvre d'une IA équitable dans la recherche, les politiques et les pratiques de santé publique.

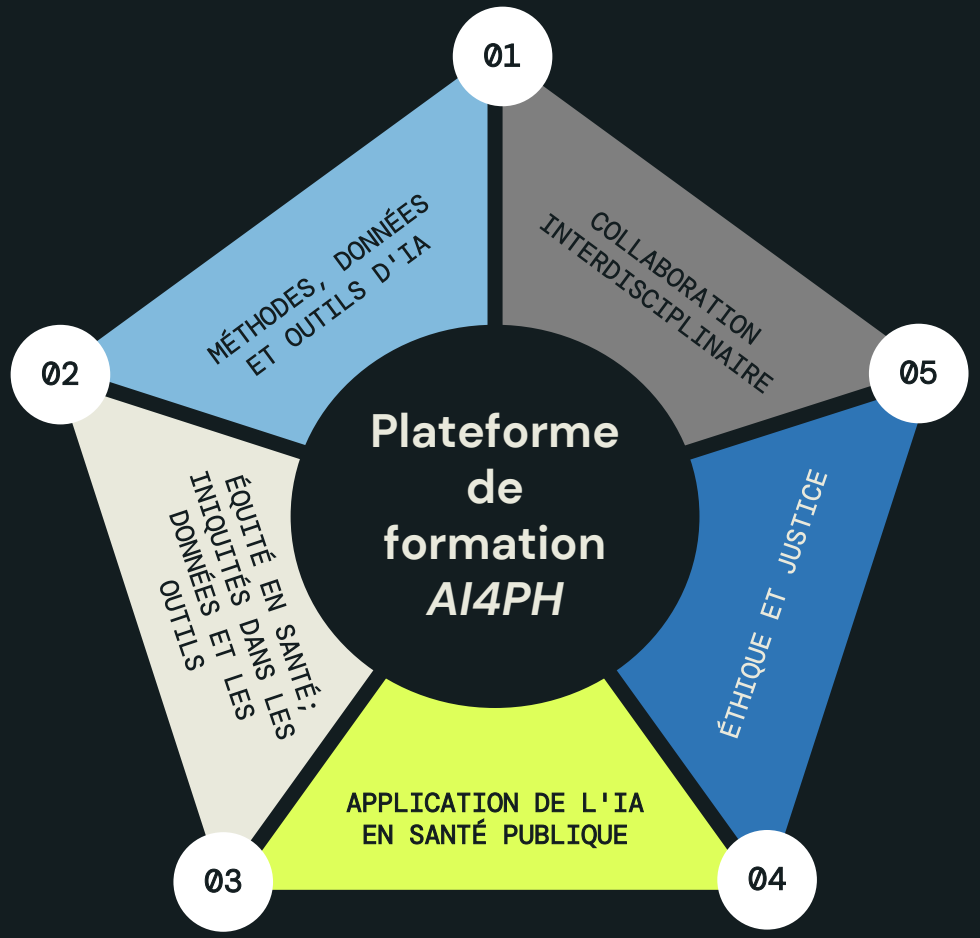
Notre plateforme mettra l'accent sur le développement des compétences et le renforcement des capacités pour appliquer les innovations en matière d'IA et d'apprentissage automatique aux défis de santé publique les plus urgents du Canada, en tenant compte des déterminants sociaux, culturels et environnementaux de la santé.

CIHR IRSC Canadian Institutes of Health Research Instituts de recherche en santé du Canada

AI4PH reçoit le soutien des Instituts de recherche en santé du Canada (IRSC)

Unclassified / Non classifié

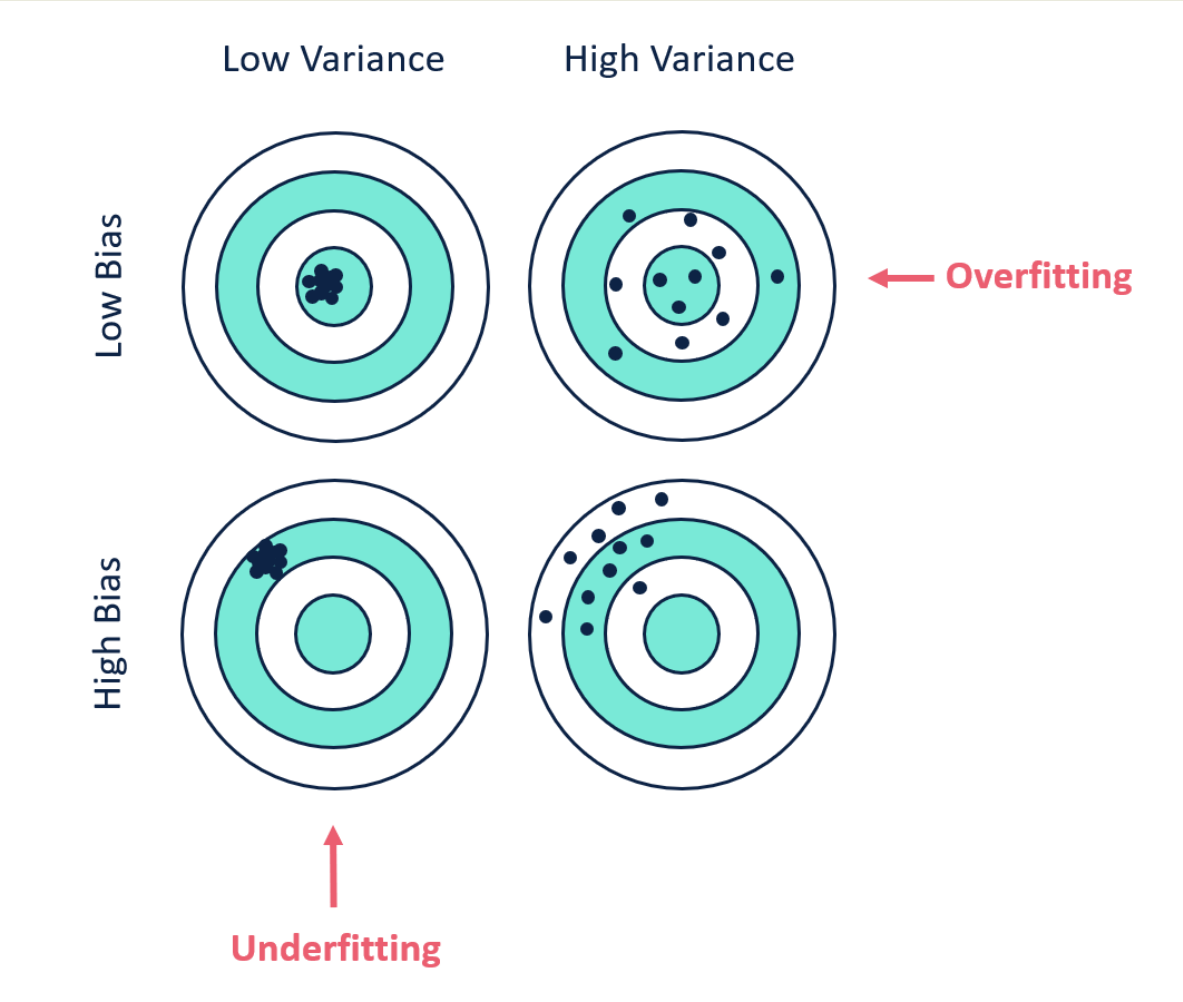
PLATEFORME DE FORMATION DE *AI FOR PUBLIC HEALTH (AI4PH)*



- **Formations courtes**
- **Bourses**
- **Stages**
- **Développement du sentiment d'appartenance**
- **Ateliers**

*English only content / Contenu en anglais seulement

Compromis entre biais et variance



- Un modèle avec une variance élevée signifie que les prédictions sont sensibles à l'échantillon d'apprentissage et peuvent refléter du bruit
- On parle de biais quand un modèle est trop simple et se trompe à plusieurs reprises
- On parle de compromis lorsqu'il faut optimiser le modèle pour trouver le bon équilibre

p. ex. un modèle plus simple aura une variance plus faible, mais s'il est trop simple, il aura plus de biais

Images : https://www.javatpoint.com/bias-and-variance-in-machine-learning

Besoin d'équivalents terminologiques?

Terme informatique	Terme épidémiologique
Fonctionnalités	Variables indépendantes
Étiquettes	Résultats
Étiquettes bruyantes	Erreur de mesure
Apprentissage	Ajustement
Exploration de données	Analyses exploratoires
Algorithme de classification	Algorithme avec variables nominales
Algorithme de régression	Algorithme avec variables continues
Précision	Valeurs prédictives positives
Rappel	Sensibilité
Mode dimensionnel	Nombre de covariables
Données déséquilibrées	Répartition inégale des variables
Encodage « un parmi n » (*one-hot*)	Création de variables nominales

Qifang et al. *What is machine learning? A primer for the epidemiologist*. AJE, 2019.