Remarque sur l'utilisation de l'IA

La transcription anglaise a été générée à l'aide de Claude Fable v5.1 (niveau d'effort moyen) le 6 octobre 2026 par Marc-André Bélair. Veuillez consulter l'annexe A pour voir la requête. **Ce texte n'a pas encore été vérifié.** Une copie du script Python qui a servi à intégrer le fichier .SRT dans cette transcription est conservée dans le dossier du projet.

Nous avons également utilisé MS Copilot pour réviser le texte en anglais à l'aide de la requête suivante : « *Write A version that stays very close to the original speaker's style. Do not change the underlying message.* ».

Le texte mis à jour a été comparé à l'original pour confirmer que le message et le sens étaient restés les mêmes.

L'IA en santé publique — Session 1 : Introduction à l'IA en santé publique (transcription)

Marc-André Bélair :

Bonjour tout le monde, quel que soit le moment de la journée chez vous selon l'endroit où vous vous trouvez aujourd'hui. Bienvenue à cette première session de la série de séminaires sur l'intelligence artificielle en santé publique, soit « Introduction à l'IA en santé publique ». Mon nom est Marc-André Bélair, et je suis épidémiologiste et formateur au sein de l'Unité d'apprentissage et de perfectionnement de la Direction générale de la réglementation, des opérations et de la gestion des urgences à l'Agence de la santé publique du Canada. Je me joins à vous depuis Vancouver, en Colombie-Britannique, et j'ai l'honneur d'être votre animateur aujourd'hui. Cette session a été présentée en anglais le 23 septembre. Nous avons traduit les diapositives ainsi que la transcription et avons procédé à un réenregistrement en français.

Avant de commencer, je voulais prendre un petit moment pour reconnaître que j'ai l'honneur et le privilège de vivre, travailler et jouer sur les terres et les eaux des territoires traditionnels non cédés des nations xʷməθkʷəy̓əm (Musqueam), Sḵwx̱wú 7mesh (Squamish) and səlilwətaɬ (Tsleil-Waututh). En tant qu'occupant et colonisateur, je leur suis reconnaissant d'avoir pris soin de cette magnifique région. Alors que nous explorons le sujet de l'intelligence artificielle dans le domaine de la santé publique, je vous encourage à réfléchir aux biais qui peuvent se trouver dans les données de formation que nous utilisons pour les modèles d'IA. De plus, alors que nous approchons de la Journée nationale de la vérité et de la réconciliation, qui aura lieu la semaine prochaine, prenez un moment pour vous intéresser et réfléchir aux peuples autochtones de ce territoire, à leurs récits et à leur histoire.

Je vais simplement vous donner un aperçu de ce qui vous attend pour la session d'aujourd'hui. D'abord, quand j'aurai fini de parler, nous aurons droit à quelques mots

d'introduction de la part de Zoe Kahn, la directrice générale du Centre des mesures d'urgence de la Direction générale de la réglementation, des opérations et de la gestion des urgences ici à l'Agence de la santé publique du Canada. Ensuite, je passerai en revue les résultats du sondage, puis nous donnerons la parole à notre invitée principale, la Dre Laura Rosella, qui nous présentera les concepts clés de l'IA en santé publique. Après la présentation de la Dre Rosella, j'aurai le plaisir d'animer une session de questions et réponses. J'aimerais inviter Zoe Kahn, notre directrice générale ici au Centre des mesures d'urgence, à faire quelques remarques introductives pour lancer la session d'aujourd'hui. Zoe, la parole est à vous.

Zoe Kahn :

Merci, Marc-André. Bonjour tout le monde, quel que soit le moment de la journée chez vous selon l'endroit où vous vous trouvez aujourd'hui. Merci d'être parmi nous. J'aimerais commencer par reconnaître qu'ici, à Ottawa, j'ai l'honneur de me trouver sur le territoire traditionnel non cédé du peuple algonquin anichinaabé, qui vit sur ces terres depuis des temps immémoriaux. Au nom de l'Agence de la santé publique du Canada et du Bureau des services de terrain et de la formation en santé publique, je suis très heureuse de vous accueillir à cette nouvelle série de séminaires intitulée « Du projet pilote à la pratique : les applications de l'IA en santé publique », que nous présentons en partenariat avec *AI4PH* et le Centre d'IA de l'ASPC. À l'ASPC, nous avons déjà organisé des conférences sur l'IA, mais ce qui distingue cette série, c'est son approche pratique. Il s'agit de présenter des applications concrètes de l'IA et d'utiliser des exemples pour expliquer ce qu'il faut pour mettre en œuvre l'IA dans des contextes réels de santé publique aujourd'hui. Il y a donc deux objectifs à cette série. D'abord, il s'agit de renforcer la compréhension et l'enthousiasme face à la façon dont l'IA peut aider à relever des défis complexes en matière de santé publique. Ensuite, il s'agit de préparer les cadres et les praticiens et praticiennes à mettre en place des projets d'IA qui mènent au succès au sein de leurs propres organisations.

D'après les sondages menés auprès des épidémiologistes de terrain sur leur utilisation de l'IA au travail, la formation qu'ils souhaitent le plus avoir est une formation pratique. Quels sont les cas d'utilisation de l'IA en santé publique? Quels outils d'IA sont les plus adaptés à des tâches précises? Ils veulent comprendre comment utiliser cette technologie de manière éthique et apprendre des autres les pièges, les erreurs, les avantages, les inconvénients, les choses à faire et les choses à éviter lorsqu'on applique l'IA à un contexte de santé publique.

Ainsi, appuyée par ces commentaires, notre série est conçue pour passer progressivement de la théorie à la pratique, de la compréhension des concepts et des méthodes de base aux témoignages d'équipes ayant mis en œuvre des projets d'IA sur le terrain. Nous savons que beaucoup de personnes, au-delà du personnel de terrain, ont les mêmes besoins de formation. Cette série s'adresse donc à un public large et averti, notamment aux professionnels et aux professionnelles de la santé publique à tous les niveaux du gouvernement, au personnel de terrain de l'ASPC au sein des différents systèmes provinciaux et territoriaux, ainsi qu'aux stagiaires, aux chercheurs et aux chercheuses, et aux collaborateurs et aux collaboratrices qui façonnent la prochaine génération d'innovations en santé publique. Grâce à cette approche « du projet pilote à la pratique »,

au cours des prochains mois, vous découvrirez des méthodes d'IA présentées par des spécialistes, de véritables projets pilotes développés dans le cadre du programme de stage d'AI4PH, qui illustrent tout le parcours, de l'identification du problème à la conception et à la mise en œuvre de solutions, et surtout, les contraintes du monde réel et les leçons tirées. Cela comprend notamment les défis liés aux données, à la gouvernance, à la confidentialité, à l'éthique et à la préparation organisationnelle.

Vous entendrez également dire que l'IA doit être utilisée de manière éthique, de façon à réduire les disparités existantes en matière de santé plutôt qu'à les aggraver. Les thèmes de l'équité et de l'utilisation responsable seront abordés tout au long de chaque session, en tant que considérations fondamentales pour l'application de l'IA dans le domaine de la santé publique.

Pour le BSTFSP et *AI4PH*, notre collaboration va au-delà d'une simple série de conférences. Nos deux groupes ont pour mission de renforcer les capacités de l'ensemble du personnel de santé publique, et nous espérons poursuivre notre partenariat au-delà de cette série de conférences pour vous proposer de futures occasions de formation. Grâce à cette série, nous souhaitons également vous aider à créer des liens avec les autres pour trouver des occasions concrètes de collaborer. Ainsi, au cours de chaque session, un lien vers Slido sera publié, lequel vous permettra d'envoyer un message aux autres personnes participant au webinaire, par exemple, pour trouver quelqu'un ayant l'expertise nécessaire pour résoudre un problème d'IA auquel vous êtes confronté, ou pour trouver des gens avec qui collaborer pour les projets sur lesquels vous travaillez, ou encore pour signaler d'autres occasions de collaboration, où vous pourrez travailler ensemble et apprendre les uns des autres. À la fin de chaque session, nous ferons circuler ces messages aux personnes inscrites pour que vous puissiez entrer en contact avec les autres et créer des liens comme bon vous semble. Pour conclure, je tiens à vous remercier de participer à cette discussion.

Nous espérons que cette série vous donnera l'inspiration et les conseils pratiques dont vous avez besoin pour intégrer l'IA dans votre travail, de manière à renforcer les pratiques de santé publique partout au Canada. Sur ce, je redonne la parole à Marc-André pour qu'il poursuive la rencontre. Merci beaucoup.

Marc-André Bélair :

Merci beaucoup, Zoe. Merci d'avoir pris le temps de nous mettre dans le bain pour la session d'aujourd'hui; c'est très apprécié. Donc, sans plus attendre, jetons un coup d'œil à certains de ces résultats. Nous voulons apprendre à vous connaître; voyons voir qui vous êtes et qui est avec nous aujourd'hui. Sur 268 personnes ayant répondu au sondage, 40 % sont des épidémiologistes, 24 % ont répondu « autre ».

Nous avons des analystes de données, des analystes de programme et de politique, des médecins hygiénistes, des infirmières et des infirmiers de la santé publique et des gens du domaine de la gestion des urgences et de la préparation. C'est donc un échantillon de personnes très varié, ce qui nous fait plaisir à voir. Ces pourcentages changeront un peu, car certaines personnes sont encore en train de répondre; c'est bien. En ce qui concerne votre rôle, plusieurs d'entre vous êtes des chercheurs, des chercheuses, des scientifiques, des universitaires ou encore des conseillers principaux et des conseillères principales. Il y a

également des praticiens et des praticiennes de première ligne avec nous, ainsi que quelques étudiants, étudiantes et décisionnaires. C'est vraiment un super mélange.

Vous êtes 70 % à avoir indiqué ne pas être chercheur ou chercheuse en début de carrière, mais 30 % le sont, ce qui est fascinant.

Ensuite, 93 % de notre public travaille principalement au Canada, contre 7 % à l'extérieur du Canada. Ici, il s'agissait d'une réponse libre dans une zone de texte, donc nous regarderons les réponses plus tard de manière individuelle. Il y avait aussi cette question à savoir quel énoncé représentait le mieux votre utilisation actuelle de l'IA. Parmi vous, 42 % disent utiliser l'IA à l'occasion pour des taches précises, 28 % disent l'utiliser régulièrement pour améliorer leur travail, 15 % l'explorent, 7 % disent qu'elle est intégrée à certains de leurs processus récurrents, 5 % disent ne pas l'utiliser et 3 % disent être à la tête d'une initiative visant l'adoption de l'IA au sein de leur équipe ou organisation. Vous connaissez donc assez bien les outils disponibles, c'est super.

Parmi vous, 48 % se sentent peu confiants quant à leur compréhension de l'IA et de la façon dont on peut l'utiliser en santé publique; c'est bien que vous soyez ici aujourd'hui. Ensuite, 25 % se sentent confiants et 20 % ne se sentent pas confiants. C'est une bonne répartition. Et par rapport au plus grand obstacle relatif à l'utilisation de l'IA dans votre travail, vous êtes 63 % à avoir signalé des craintes quant aux politiques, à la confidentialité et à la gouvernance ainsi que le manque de confiance envers les résultats de l'IA. Il y a également le manque de compétences et de formation, la préparation organisationnelle, la qualité des données ou l'accès à celles-ci, puis 5 % des personnes ont répondu « Autre ». On peut donc voir qu'il y a plusieurs inquiétudes, ce qui est très utile pour nous. Maintenant, je vais arrêter mon partage d'écran.

Merci beaucoup de m'avoir donné l'occasion de revoir ceci. J'ai assez parlé.

Vous êtes ici aujourd'hui pour notre grande vedette. J'ai donc l'honneur et le privilège de vous présenter la Dre Laura Rosella, qui dirigera la conférence d'aujourd'hui. La Dre Laura Rosella est professeure à la *Dalla Lana School of Public Health* de la *University of Toronto* et titulaire de la Chaire de recherche du Canada en analytique de la santé des populations. Ses recherches portent sur l'analyse des données relatives à la santé des populations, la modélisation prédictive et l'utilisation de la science des données et de l'apprentissage automatique pour améliorer la santé publique et la prise de décision en matière de planification des systèmes de santé. La Dre Rosella dirige de nombreuses collaborations avec des gouvernements, des organisations de santé publique et des systèmes de santé partout au Canada pour transformer les analyses avancées en solutions concrètes face aux défis réels de la santé publique. Elle est directrice scientifique et titulaire de la chaire de recherche de la famille Stephen en santé communautaire à la *Institute for Better Health* de *Trillium Health Partners*, et fondatrice du *Population Health Analytics Laboratory* à la *Dalla Lana School of Public Health*. Elle a été désignée chercheuse principale pour *Artificial Intelligence for Public Health* ou *AI4PH*, une plateforme nationale de formation à la recherche en santé financée par les Instituts de recherche en santé au Canada. Dre Rosella, merci de vous joindre à nous aujourd'hui. La parole est à vous.

Dre Laura Rosella :

Merci beaucoup, Marc-André, pour cette belle présentation, et pour l'invitation d'aujourd'hui. Je suis vraiment emballée par cette série de conférences, à commencer par celle d'aujourd'hui, et je suis très contente de voir que vous êtes là en grand nombre; je reconnais d'ailleurs quelques noms de collègues dans le clavardage. C'est vraiment super d'être ici avec vous. Avant de commencer, je veux juste être certaine que vous voyez mon écran.

Marc-André Bélair :

(Marc-André : Je le vois très bien, oui.)

Dre Laura Rosella :

Je me joins à vous depuis la *University of Toronto*. J'aimerais reconnaître le territoire sur lequel se trouve la *University of Toronto*. Depuis des milliers d'années, c'est le territoire traditionnel des Wendat, des Sénécas et des Mississaugas de Credit. Encore aujourd'hui, il s'agit du foyer de nombreux peuples autochtones venus de l'Île de la Tortue.

Nous sommes reconnaissants de pouvoir travailler sur ce territoire. Notre sujet d'aujourd'hui est l'IA en santé publique. Je pense que c'est le moment idéal pour nous demander ce que cela signifie dans le contexte des données : à qui appartiennent-elles, quelles connaissances guident notre travail, qui a autorité sur ces données et comment les communautés, et en particulier les communautés autochtones, participent-elles aux discussions que nous avons en ce moment? Je tiens aussi à souligner qu'il y a un nombre impressionnant d'érudits et de leaders autochtones qui animent vraiment le débat sur l'IA d'un point de vue autochtone, et c'est une super occasion pour tout le monde de s'impliquer et d'écouter ce qui se passe dans ces espaces.

Aujourd'hui, j'ai le privilège d'être la première à prendre la parole dans le cadre de cette série, je vais donc poser quelques bases et vous mettre en contexte pour le reste de la série. Je vais parler des concepts fondamentaux brièvement, parce qu'on ne peut pas vraiment aborder les aspects pratiques sans d'abord comprendre conceptuellement ce qui se cache derrière. Ensuite, le plus important, je vais discuter de la façon dont l'IA est utilisée et dont elle peut apporter une valeur ajoutée. Nous aurons ensuite une discussion plus large sur les répercussions. Je tiens à préciser qu'on pourrait passer une heure et demie sur n'importe lequel de ces aspects, soit la santé, l'équité, les conséquences environnementales, mais je n'aborderai que quelques-uns d'entre eux. Et enfin, qu'est-ce qu'il faut pour mettre le tout en œuvre? Je pense que c'est l'esprit même de cette série : qu'est-ce que cela signifie pour la santé publique, et comment avance-t-on?

Je commence toujours une discussion sur l'IA et la santé publique avec cette diapositive, car je pense que, lorsqu'on parle d'une technologie transformatrice comme l'IA, on peut être tenté de se concentrer sur la technologie elle-même, mais on devrait toujours concentrer notre attention sur les répercussions sur la santé publique. Pour faire simple, comment savoir si les avancées en matière d'IA ont des répercussions positives sur la santé publique? Si c'est le cas, il y aura moins de gens qui développent des maladies ou qui souffrent de problèmes de santé évitables. Nous ferons la promotion de la santé dans nos environnements et nos communautés, et nous réduirons les iniquités en santé. Si ce n'est

pas le cas, nous devons nous demander pourquoi. Nous devons concentrer notre attention là-dessus. Il peut y avoir beaucoup de distractions et de bruit. Je pense que l'une des tâches les plus importantes pour quiconque s'intéresse à ce sujet est justement de réduire ce bruit et de concentrer nos efforts sur ce qui nous permettra d'atteindre ces résultats positifs.

[Cette diapositive comprend un diagramme montrant les fonctions essentielles de santé publique, y compris la promotion de la santé, la prévention des maladies, la protection, la surveillance, la préparation aux urgences, les politiques, le personnel et les connaissances.] Il faut également prendre du recul. Si l'on y réfléchit dans une perspective plus large, quelle place occupe l'IA dans ces fonctions essentielles de la santé publique?

Certaines de ces fonctions concernent des domaines d'activités précis, de la promotion de la santé à la préparation aux urgences, mais elles touchent aussi notre travail sur les politiques et les programmes, notre personnel et notre synthèse des données et des connaissances. Il y a donc de nombreuses occasions où l'IA peut apporter sa contribution. Il y a aussi des cas où on ne devrait peut-être pas se précipiter vers l'IA, et nous parlerons de comment faire la différence; nous devons comprendre toutes les répercussions possibles. J'espère donc qu'après aujourd'hui, vous aurez une bonne idée de tout ça.

[Cette diapositive comprend une ligne du temps illustrant les développements majeurs de l'intelligence artificielle des années 50 aux années 90, notamment le test de Turing, ELIZA, les systèmes experts, les hivers de l'IA et Deep Blue.] Avant d'aborder certains des principes fondamentaux, je crois que l'histoire est importante alors qu'on entre dans l'ère de ce qui sera l'une des technologies les plus transformatrices qu'on ne verra jamais. Tout d'abord, l'IA n'est pas un concept nouveau. Ses fondements méthodologiques remontent aux années 50. Alan Turing a écrit son article fondateur « Les machines peuvent-elles penser? ». Il y a eu une réunion très célèbre à ce propos, ainsi que quelques premiers développements. Au début, beaucoup d'optimisme régnait.

Encore une fois, ça vous dit peut-être quelque chose. Les discours étaient : « Cette technologie va régler beaucoup de problèmes », « Cette technologie nous permettra de faire plein de choses que nous ne pouvions pas faire avant », etc. Il y avait déjà des applications dans le domaine de la santé dès les années 70, mais ce n'était pas les meilleures. Il y avait une technologie MYCIN basée sur des règles, conçue pour identifier les infections bactériennes, mais celle n'était pas très performante. En fait, il y a eu d'autres projets qui n'étaient pas terribles non plus. On a connu deux de ce qu'on appelle des « hivers de l'IA ».

Il est important de comprendre ce qu'est un « hiver ». Après une période où l'on a connu beaucoup d'enthousiasme autour d'un sujet, il peut y avoir une période où l'on connaît beaucoup de critiques sur le potentiel réel d'une technologie et peut-être quelques échecs. Après donc une période de grande excitation et d'activité, le domaine a connu une longue période calme. Sa renaissance a eu lieu dans les années 90, et elle était vraiment liée aux jeux. Nous avons commencé à voir des ordinateurs gagner des jeux et des tournois d'échecs, et cela a relancé l'enthousiasme.

[Cette diapositive comprend une ligne du temps illustrant les développements majeurs de l'intelligence artificielle des années 2000 à aujourd'hui, notamment ImageNet, les avancées de l'apprentissage profond, AlphaGo, ChatGPT, AlphaFold et l'IA multimodale.] L'IA dont on

parle aujourd'hui s'est largement définie après 2010, et je vous expliquerai pourquoi, mais on a vu un regain d'intérêt. La capacité de calcul a augmenté, tout comme le volume de données numériques. En 2010, nous avons connu des avancées majeures dans le domaine de l'apprentissage profond. C'est pour cette raison que Jeff Hinton a remporté le prix Nobel, soit pour ces avancées en matière d'apprentissage profond. Le plus gros progrès concernait les images. Il y a eu plein de compétitions pour voir si l'on pouvait identifier et classer correctement une image à l'aide d'un ordinateur. Cela a donné lieu à plein d'avancées techniques. En 2020, Alpha Fold est apparu, déterminant la structure des protéines, ce qui a vraiment transformé la science fondamentale et l'IA. Nous avons eu ChatGPT, qui a transformé la façon dont les gens interagissent avec l'IA, et maintenant, on parle d'agents, une intelligence autonome et généralisée dont je vais parler dans un instant.

Encore une fois, si je vous présente cet historique, c'est pour vous montrer que ce n'est pas nouveau. Ça existe depuis longtemps, et il y a eu plusieurs vagues d'activité. Les avancées techniques après 2010 ont vraiment permis de mettre au point la technologie que l'on voit aujourd'hui.

[Cette diapositive présente un graphique d'un rapport de Stanford sur l'indice de l'IA montrant les améliorations rapides des systèmes d'IA dans plusieurs tâches de référence, de nombreux systèmes atteignant désormais, voire dépassant, les performances humaines.] Chaque année, j'attends ce rapport de référence publié par Stanford par rapport à l'indice de l'IA avec impatience. Jusqu'à présent, les progrès de l'IA ont surtout consisté à battre des records, notamment ceux établis par les humains. Pensez par exemple à la reconnaissance de l'écriture manuscrite. Dans quelle mesure un humain est-il capable de reconnaître l'écriture manuscrite par rapport à un ordinateur?

On peut voir la référence humaine, mais je veux que vous regardiez d'abord la forme de ces lignes. Vous voyez qu'elles évoluent comme ça. La progression de cette référence est extrêmement rapide. Ça a beaucoup changé depuis les débuts. J'ai examiné les graphiques des années 90 dans ce domaine, où il fallait 10 ans aux ordinateurs pour atteindre le niveau des humains en matière de reconnaissance de l'écriture manuscrite. Ce rythme est en fait vertigineux. C'est très important que l'on garde cela à l'esprit. Ensuite, les tâches sont bien plus complexes.

J'en parlerai plus tard. On ne se contente plus de reconnaître des schémas et de faire des prédictions. L'IA est l'équivalent des mathématiques et du raisonnement de niveau doctorat, ce que les gens pensaient que l'IA ne pourrait jamais faire, le raisonnement étant l'une des tâches humaines les plus complexes, voire la plus complexe. On constate déjà, dès 2026, de nombreuses capacités scientifiques et de raisonnement de niveau doctorat. Ces capacités existent donc bien, et elles gagnent en rapidité.

[Diagramme imbriqué illustrant les liens entre l'intelligence artificielle, l'apprentissage automatique et l'apprentissage profond, l'apprentissage profond étant un sous-ensemble de l'apprentissage automatique, et l'apprentissage automatique un sous-ensemble de l'IA.] Encore un petit mot sur les principes fondamentaux : j'ai déjà mentionné ces termes. Je n'aime pas donner de définition précise de l'IA, car, en quelque sorte, elle n'en a pas qu'une seule. Il s'agit vraiment d'un terme générique. Lorsque je parle d'IA, je fais référence

à une vaste catégorie de méthodes qui consistent à utiliser des approches informatiques pour faire quelque chose que l'on considérait normalement comme nécessitant l'intelligence humaine. Encore une fois, ça peut aller du très simple au très complexe.

L'apprentissage automatique consiste en fait à identifier des schémas à partir de données, et l'apprentissage profond, comme vous l'avez vu dans l'aperçu historique, a accéléré ce processus pour qu'on puisse exploiter des données complexes. Je ne vous ferai pas une présentation technique là-dessus, mais l'apprentissage profond consiste à décomposer des données complexes en éléments plus petits pour que l'on puisse les traiter par ordinateur et résoudre des problèmes qui étaient auparavant insolubles. C'est pour cette raison qu'on le considérait comme une percée. Prenons l'exemple des images. C'était très difficile pour un ordinateur d'analyser une image. Quand on la transforme en réseau neuronal convolutif et qu'on la décompose, c'est possible. Je simplifie le tout, évidemment. Ça ne concerne pas seulement les images, mais aussi le texte, le son et le langage. C'est la technologie sous-jacente à beaucoup de nos grands modèles de langage, comme GPT et Claude.

Je dois parler un instant de l'IA générative, juste pour que tout le monde comprenne bien de quoi il s'agit, parce que c'est probablement l'utilisation la plus populaire de l'IA, pas forcément au travail, même si c'est possible, mais la plupart des gens interagissent avec une forme d'IA tous les jours. La différence avec l'IA générative est qu'il s'agit de tirer des schémas à partir de données existantes que l'on connaît, mais de créer en réalité du nouveau contenu, un nouveau langage, ce qui correspond à la façon dont la plupart des gens utilisent ces modèles. Ça concerne aussi la musique; peut-être que vous avez déjà entendu de la musique générée par l'IA, ou encore le son et les images; peut-être que vous avez déjà travaillé avec des images générées par l'IA. Encore une fois, il s'agit d'un petit changement, d'une catégorie différente. Pour ramener le tout à notre discussion sur la santé publique, les gens s'en servent pour créer des rapports, des lignes directrices, en adaptant ces dernières à différents publics. Cela demande quand même une vérification. L'IA générative peut engendrer beaucoup de problèmes, et, encore une fois, nous aborderons certains d'entre eux lors de nos prochaines sessions plus détaillées. Il y a les hallucinations, les fabulations, etc.

Les modèles s'améliorent sans cesse, mais il y a plein de choses à prendre en compte ici.

Passons en revue quelques points techniques supplémentaires avant de passer aux exemples. Je tiens à souligner que quand vous travaillez dans ce domaine, les modèles ne sortent pas de nulle part, n'est-ce pas? Il y a des étapes vraiment très importantes. Les modèles doivent être entraînés, puis validés. Il existe des approches très rigoureuses à ce sujet. Il y a désormais des normes. Quand j'ai commencé dans ce domaine il y a neuf ans, presque 10 ans, croyez-le ou non, il n'y avait pas tellement de normes entourant l'apprentissage automatique. Maintenant, il y a beaucoup plus de normes sur ce que l'on considère comme une bonne validation et de bons tests. Ce que je veux souligner sur cette diapositive, par contre, c'est qu'il faut, selon moi, mettre fortement l'accent sur la santé publique. D'ailleurs, la santé publique dispose d'une grande expertise. Il s'agit donc de savoir comment évaluer ces outils dans des conditions réelles pour s'assurer qu'elles améliorent bien les résultats. Ces trois résultats dont j'ai parlé au début : avons-nous de meilleurs résultats en matière de santé publique? Comment suit-on l'évolution des modèles

au fil du temps? C'est vraiment plus compliqué que d'évaluer d'une intervention statique qui ne change pas et de simplement mesurer les résultats. Il faut en fait un suivi continu. Je vais y revenir à la fin, mais l'une des lacunes est de trouver comment intégrer l'évaluation et le suivi en conditions réelles au sein de nos institutions de santé publique et dans nos contextes. C'est l'une des choses les plus difficiles que l'on doive faire en ce moment.

Ensuite, pour ceux et celles qui se disent que c'est encore trop abstrait pour eux, voici quelques exemples de la façon dont on utilise le tout. Je vais aborder des exemples concrets dans quelques minutes, mais on a par exemple le regroupement, qui consiste à identifier les regroupements dans les données, et qui est très utilisé dans le domaine de la surveillance. Nous avons, comme je l'ai déjà mentionné, l'apprentissage profond appliqué aux données complexes. Un domaine que je tiens à souligner, c'est celui des images environnementales. Par exemple, chez *AI4PH*, nous avons organisé un atelier d'été sur l'IA et la création d'images. Nous ne parlions pas d'imagerie comme les radiographies thoraciques. Nous parlions de l'imagerie de l'environnement bâti, par exemple, ou de l'imagerie environnementale. Je vous donnerai l'un de ces exemples un peu plus tard. Bien sûr, il y a également les grands modèles de langage, qui constituent le fondement de l'IA générative, ainsi que ce qu'on appelle le plus en plus les modèles multimodaux.

Essentiellement, cela signifie qu'il ne s'agit pas seulement de prendre des données d'une seule source de données, mais d'en prendre de plusieurs flux de données à la fois. Techniquement, c'est un peu plus difficile, mais si vous pensez à la diapositive de capacité, nous sommes maintenant au point où cela devient moins un obstacle technique et sera la référence de toute l'IA moderne. Tout cela reposera sur des données multimodales.

Et si l'on songe un peu à l'avenir, juste pour vous faire réfléchir, on pourrait en parler longuement : où allons-nous? Encore une fois, c'est pour cette raison que j'aime présenter ces tendances de référence dès le début. Nous nous orientons vers ce qu'on appelle les « modèles fondateurs », qui s'éloignent des modèles très spécifiques à des tâches précises pour s'orienter vers des modèles plus généraux. Encore une fois, ces modèles sont à la base d'une grande partie de la technologie de l'IAG. On les appelle des modèles transformeurs préentraînés. C'est ce que signifie l'acronyme « GPT » en anglais. Ils peuvent effectuer de nombreuses tâches différentes en temps réel. Nous nous éloignons de l'utilisation d'un seul modèle une seule fois. En fait, comment le modèle continue-t-il à recevoir de l'information et à se réentraîner? Parce que le modèle devrait tirer parti de son expérience d'utilisation pratique pour s'améliorer.

L'IA multimodale, j'en ai déjà parlé. L'IA agentive, qui représente un changement majeur et constitue actuellement une priorité, consiste à passer d'une utilisation d'un modèle unique pour une tâche donnée à un système capable de planifier et de mettre en œuvre des outils. Ainsi, il effectue une prédiction, exécute une action et réalise l'ensemble du processus, du début à la fin. Nous en reparlerons. Enfin, le dernier point, qui fait l'objet de nombreux débats philosophiques quant à savoir si et quand cela se produira, est le concept de notre intelligence généralisée couvrant de nombreuses tâches.

Avant de passer aux exemples, je vous ai donc présenté en quelque sorte les bases. Il s'agit de la partie la plus technique de cette présentation. Si vous avez trouvé que c'était

beaucoup, c'est probablement le cas, mais nous avons posé les bases techniques, pour aujourd'hui et pour toute la série de conférences. C'était pour que vous compreniez un peu mieux que l'IA n'est pas simplement un concept. Ce sont des choses concrètes, et vous savez de quoi il s'agit. À ce niveau conceptuel très simple, j'aimerais que vous considériez l'IA comme accomplissant l'une de ces tâches. Si, au lieu de vous concentrer sur ce qu'est le modèle sous-jacent, vous vous concentrez davantage sur ce à quoi sert ce modèle, le tout devient beaucoup plus concret. Alors, « volet descriptif », « volet prédictif », « inférence causale » et « volet prescriptif » : qu'est-ce que cela signifie? Souvent, l'IA sert simplement à vous dire ce qui se passe.

Elle peut également vous dire ce qui va se passer.

Dans des cas précis, elle peut vous dire ce qui pourrait changer le cours des choses, mais c'est beaucoup plus difficile. Finalement, dans les cas qui relèvent du volet prescriptif, il s'agit en fait de remplir le processus décisionnel, ce qui, encore une fois, est très rare. Nous avons donc beaucoup de cas dans les deux premières catégories, très peu dans les deux dernières, ce qui semble approprié, étant donné que nous souhaitons avoir un contrôle humain complet sur la plupart des applications envisagées en santé publique. Il y a donc beaucoup de surveillance. Dans le volet prédictif, vous pourriez utiliser l'IA pour établir les priorités et faire la planification. Dans la catégorie causale, il s'agit d'interventions ou de politiques, tandis que les recommandations et les systèmes automatisés sont, encore une fois, les moins fréquents à l'heure actuelle.

[Cette diapositive présente un schéma illustrant les domaines dans lesquels l'IA peut soutenir les organisations de santé publique : les services destinés au public, l'analyse et l'aide à la décision ainsi que les opérations à l'interne.] Passons donc à quelques exemples, car cela vous donnera une bonne idée de la manière dont l'IA est réellement utilisée et de ce que nous pouvons tirer de ces exemples. Il s'agit uniquement d'exemples publics au sujet desquels vous pouvez lire davantage, c'est d'ailleurs pour cette raison que je les ai choisis, et je vais vous en présenter quelques-uns.

Avant de vous donner les exemples, je tiens à rappeler, comme je l'ai dit au début, qu'il y a beaucoup de bruit autour de ce sujet. L'IA peut être utilisée de plusieurs façons différentes, et il existe un écart important lorsque l'on parle de l'IA utilisée pour les opérations à l'interne. Je suis certaine que vos organisations ou vos institutions, ainsi que vous-mêmes, envisagent d'appliquer l'IA aux flux de travail opérationnels, au traitement des documents et aux tâches administratives. Il y a ensuite le volet lié à la santé publique, où l'on pense à la surveillance, peut-être à l'automatisation de la synthèse des données probantes, à l'évaluation des risques, en utilisant des méthodes plus complexes. Et puis, il y a cet aspect orienté vers le public, que l'on observe dans certaines organisations, comme les robots conversationnels, qui fournissent de l'information à la communauté. Évidemment, ces contextes présentent des profils de risque complètement différents, mais je tiens à souligner que l'IA est présente dans tous ces contextes, et que des exemples peuvent s'y appliquer. Et si vous pensez à l'IA d'un point de vue stratégique au sein d'un organisme de santé publique, vous pensez probablement à des exemples concrets dans chacun de ces contextes.

Permettez-moi donc de vous présenter le premier exemple. Il s'agit en fait d'étudier les maladies d'origine alimentaire et de transformer le texte libre issu des antécédents alimentaires en une analyse standardisée des éclosions d'origine alimentaire. C'est donc un exemple très concret. Un modèle de langage classe ces catégories alimentaires non structurées en catégories permettant théoriquement d'identifier rapidement les informations d'exposition à grande échelle. Il y a probablement beaucoup de gens ici qui font ce genre de travail, et le temps est un facteur crucial. Lorsqu'une éclosion survient, nous voulons connaître le plus tôt possible la source de l'éclosion ou la cerner davantage pour mener une enquête. Une enquête sera toujours nécessaire. Cela permet d'accomplir cette tâche beaucoup plus rapidement, en utilisant les observations sous forme de texte libre. Il s'agit d'un exemple utile.

C'est très compliqué du point de vue du modèle, mais il s'agit en réalité d'une application descriptive et de surveillance. Encore une fois, on peut s'attarder sur la complexité de la technologie, mais cet exemple décrit en fait un outil réellement utilisé à des fins de surveillance, et des mesures concrètes en découleront probablement. C'est un exemple utile, car il est facile à localiser. Il y a beaucoup de texte à organiser. La tâche de classification est très clairement définie. Il y a des gens, des responsables de la santé publique, qui utiliseront ces données et les valideront, et nous pourrons clairement évaluer si cela fonctionne ou non.

Voilà donc un bel exemple d'application de l'IA dans le domaine de la santé publique.

Voici un autre exemple, très simple cette fois-ci, mais encore très concret. Il existe un outil appelé « MedCoder », utilisé par le *CDC* pour convertir les causes de décès indiquées sur les actes de décès en codes normalisés. Tous les gens qui travaillent avec des données sur les décès, c'est mon cas depuis de nombreuses années, savent que ce processus prend beaucoup de temps. Les processus associés sont très manuels. Grâce à l'apprentissage automatique et au traitement du langage naturel, ils ont réussi à coder automatiquement près de 90 % des dossiers. Certains n'ont pas pu être codés et ont dû être traités manuellement, mais il s'agit encore une fois d'une approche descriptive. Cette méthode est souvent utilisée à des fins de surveillance et d'établissement des priorités.

Et cela peut se faire plus rapidement, ce qui permet d'accélérer toutes les étapes suivantes. L'un des exemples cités à cet égard concerne la surveillance des décès par surdose d'opioïdes : il s'agit de localiser et d'identifier plus rapidement ces cas afin d'orienter les mesures de santé publique.

Je vais vous donner encore quelques exemples très rapidement. [Cette diapositive présente des graphiques comparant un modèle d'apprentissage automatique et des approches traditionnelles pour prédire le risque d'hospitalisation lié à la COVID-19, montrant une meilleure identification des personnes à haut risque.] Il s'agit d'un exemple lié à la COVID, qui porte sur la manière dont nous pouvons estimer le risque d'hospitalisation chez les personnes dont le test de dépistage de la COVID s'est révélé positif, dans le but d'établir les priorités en matière de vaccination. Nous savons donc que nous avons établi les priorités de vaccination en nous basant sur des critères que nous savons importants, soit l'âge, les comorbidités et certains autres facteurs, mais il s'agissait de systèmes simples fondés sur

des règles. Au contraire, avec cette étude, nous avons pu démontrer qu'en utilisant les données existantes, sans avoir besoin de recueillir de nouvelles données, nous pouvons effectuer l'analyse de manière beaucoup plus précise et rapide. Cela démontre donc que nous pouvons identifier rapidement ces personnes, ce qui pourrait alimenter les processus en temps réel.

[Cette diapositive présente un exemple d'utilisation de la vision par ordinateur pour identifier des tours de refroidissement potentielles à partir d'images aériennes, dans le cadre d'enquêtes sur *Legionella*.] C'est un sujet vraiment intéressant qui a suscité beaucoup d'attention. Encore une fois, cela implique l'apprentissage profond et des méthodes de vision par ordinateur. Il s'agit donc de méthodes très sophistiquées. Je suis certaine que plusieurs d'entre vous ont été confrontés aux éclosions de légionellose qui ont touché l'ensemble du Canada ces deux dernières années. En général, une étape clé de l'investigation consiste à identifier les tours de refroidissement. Une personne travaillant sur ces éclosions au sein d'une unité locale de santé publique m'a expliqué qu'il faut repérer les tours de refroidissement, ce qui nécessite souvent de parcourir les environs en voiture ou de regarder et de se demander où se trouvent les tours de refroidissement, puis de mener une enquête approfondie. Or, avec l'IA, la vision par ordinateur permet de passer en revue les images aériennes en quelques instants pour repérer les tours, puis d'affiner la recherche. L'action de santé publique ne change donc pas beaucoup ici. Cette enquête doit tout de même être menée.

C'est simplement une approche plus précise, et on veille à ce que le tout puisse se dérouler plus rapidement et avec plus de précision, en tirant parti des avancées en matière d'apprentissage automatique. De très bonnes méthodes ouvertes et accessibles ont été mises au point pour y parvenir.

[Cette diapositive présente une série de graphiques illustrant comment l'IA classe les publications sur les réseaux sociaux selon qu'elles expriment une position positive, négative ou neutre par rapport à la vaccination.] Pour terminer, je voulais mettre de l'avant un exemple d'utilisation des modèles de langage pour analyser le discours du public. Ce travail provient de l'équipe de Laura Espinosa, qui a utilisé un modèle de langage accessible au public pour déterminer si les publications sur les réseaux sociaux reflétaient une position positive, négative ou neutre par rapport à la vaccination. Ce type d'analyse du discours est très courant en santé publique, et de nombreux progrès ont été réalisés pour permettre de prioriser les efforts. Tous ces exemples ont des limites, et nous pourrions avoir une discussion très approfondie sur les forces et les limites, mais je tenais à vous donner une idée du fait que ces modèles préentraînés offrent une façon très évolutive de procéder. Ce travail est mené depuis plusieurs années dans le domaine de la santé publique, mais une analyse des sentiments plus traditionnelle nécessiterait un important travail manuel d'étiquetage des publications.

Quels enseignements peut-on tirer de ces exemples? Je vous ai donné une série d'exemples; dans la plupart des cas, on voit l'IA structurer l'information, identifier des tendances et faciliter l'établissement des priorités, mais chacun de ces exemples devait être lié à une mesure de santé publique. Dans certains exemples, ce lien était plus évident que dans

d'autres. Je dirais qu'il est vraiment important de partir de ce point de départ. Quelles mesures de santé publique pouvons-nous mettre en œuvre?

Ou quel flux de travail pouvons-nous mettre en place? Et quelles sont les possibilités offertes par ces outils? Je n'ai pas abordé la question de l'évaluation de tous ces modèles, mais nous devrions réfléchir à ces évaluations. Ces modèles sont-ils efficaces? Sont-ils utiles? Ne le sont-ils pas? Cela va bien au-delà de la performance. Je pense que la performance a dominé la littérature dans ce domaine pendant de très nombreuses années. Aujourd'hui, ce que nous cherchons vraiment à savoir, c'est si cela change réellement les résultats.

Cela a-t-il une incidence sur la charge de travail? L'objectif, qui est idéalement de prévenir une maladie évitable, permet-il d'accomplir toutes ces tâches plus rapidement et plus efficacement? Cela permet-il de rejoindre davantage de personnes? Cela permet-il de rejoindre les bonnes personnes? C'est ainsi que nous devons aborder ces évaluations. Et bien sûr, il y a des avantages et des inconvénients à prendre en compte, qui doivent être intégrés à notre analyse.

Avant d'aborder certains de ces défis, je tiens à réitérer un point que j'ai déjà soulevé à plusieurs reprises, car je pense qu'il revêt une importance particulière pour la santé publique. Nous devons faire évoluer le débat, qui ne doit plus porter sur « Devrions-nous utiliser l'IA? » ou « Comment utiliser davantage l'IA? », mais plutôt sur quelque chose de plus concret, comme « Si l'IA m'aide à apprendre X plus rapidement, plus facilement ou avec davantage de confiance, quelle décision vais-je prendre différemment, et en quoi cela améliorera-t-il un résultat mesurable? ».

Pour les décisionnaires en santé publique, ce lien est essentiel. L'IA possède des capacités exceptionnelles, c'est pourquoi j'ai commencé par là. Elle peut générer des connaissances et traiter des informations à une échelle et à une vitesse que je n'aurais jamais imaginé atteindre en 2026. Cependant, l'objectif ne concerne pas les capacités elles-mêmes. La vraie question est de savoir ce que nous pouvons faire de cette information et si cela mène à de meilleurs résultats pour les individus, les communautés ou les systèmes de santé publique. Je veux être claire, je suis absolument convaincue qu'il existe de solides cas d'utilisation où la réponse est oui.

Pour l'instant, ce lien n'est pas aussi clair qu'il pourrait l'être. Plus vite nous y parviendrons, plus vite nous pourrons concentrer notre attention sur la valeur ajoutée plutôt que de nous contenter d'ajouter une nouvelle application d'IA simplement parce que nous le pouvons.

En passant, j'inclus les tâches administratives dans cette réflexion. Je sais que ces tâches sont extrêmement importantes pour le travail quotidien et la charge de travail des gens. Malgré cela, nous devrions nous demander quelles applications nous aident à faire progresser les objectifs de santé publique que nous cherchons à atteindre, que ce soit directement ou indirectement.

Parlons donc de certaines de ces répercussions.

Je vous ai donné un aperçu de ce qu'est l'IA, de la rapidité avec laquelle elle évolue, de la direction que prend cette technologie ainsi que quelques exemples de son utilisation actuelle en santé publique. Je voudrais maintenant aborder certaines de ces répercussions. Ce sont des sujets que nous explorerons de manière beaucoup plus approfondie tout au long de la série de conférences, mais je souhaite en présenter quelques-uns aujourd'hui.

Plus précisément, je voudrais parler des répercussions sur les données, l'équité et l'environnement. Chacun de ces sujets pourrait facilement faire l'objet d'une séance à part entière, mais, pour l'instant, nous allons brièvement passer en revue les considérations clés, puis discuter de la manière dont nous pouvons transformer cette compréhension en une approche de mise en œuvre plus réfléchie et responsable.

Le secteur de la santé publique travaille avec des données depuis des siècles; à bien des égards, ce n'est donc pas un nouveau défi pour nous. Nous avons toujours dû porter un regard critique sur la qualité des données, leur représentation, leur biais et leur adéquation réelle à l'objectif pour lequel nous les utilisons. L'IA ne crée pas ces défis, mais elle les amplifie.

Parallèlement, l'un des plus grands avantages de l'IA est qu'elle nous permet de travailler avec des types et des volumes de données auparavant inaccessibles. Prenons l'exemple de l'utilisation de l'imagerie pour identifier les tours de refroidissement. C'est quelque chose que nous ne pouvions tout simplement pas faire à cette échelle et à cette vitesse auparavant.

Cependant, à mesure que ces capacités s'étendent, il est important que nous restions critiques et réfléchis quant à la manière dont les données sont utilisées. C'est là que, selon moi, la santé publique a un rôle unique à jouer. Nous disposons d'une expertise approfondie pour poser les bonnes questions concernant les données : qui est représenté? Qui ne l'est pas? Les informations sont-elles fiables? Quelles conclusions peut-on en tirer?

C'est pour cette raison que, lorsque les organisations conçoivent et mettent en œuvre des applications d'IA, la santé publique doit avoir son mot à dire. Qui est le mieux placé pour évaluer la qualité des données, leur représentativité et leur utilisation appropriée? À ces égards, je crois que la santé publique est mieux placée que presque tous les autres secteurs pour aider à répondre à ces questions.

Voici quelques-unes des questions que nous devrions nous poser : l'utilisation de ces données permettra-t-elle d'obtenir de nouvelles informations?

Il s'agit d'une question importante, car certaines données sont fascinantes, et il est impressionnant que l'IA puisse les traiter, mais nous devons tout de même nous demander si elles nous apprennent réellement quelque chose que nous ignorions auparavant. Génère-t-elle quelque chose de nouveau? Nous aident-elles à prendre une meilleure décision?

Une autre question cruciale, sur laquelle je reviendrai, car elle est particulièrement pertinente lorsque nous parlons de répercussions environnementales, est la suivante : qui ces données représentent-elles?

Le biais de sélection est un problème important. Quiconque travaille en santé publique, en particulier les épidémiologistes, a l'habitude de se poser cette question presque instinctivement. Même dans le cas d'une estimation de prévalence, qui semble pourtant simple, nous savons que les données ne représentent pas tout le monde. Nous devons donc comprendre qui est représenté, qui ne l'est pas, et comment ces lacunes pourraient influencer les décisions en matière de santé publique que nous prenons.

Parfois, la valeur ne réside pas dans le fait que l'information soit nouvelle. Parfois, la valeur réside dans le fait que nous puissions l'obtenir plus rapidement. C'est une bonne chose, mais seulement si cette rapidité mène à un meilleur résultat.

Si l'IA se contente de produire de l'information rapidement, mais que celle-ci alimente ensuite un processus qui continue d'avancer à pas de tortue, nous risquons de ne pas en tirer grand-chose. Dans certains cas, nous pourrions même créer de nouveaux obstacles. Cependant, si un accès plus rapide à l'information nous permet d'agir plus tôt, d'atteindre davantage de personnes, de cibler les interventions plus efficacement ou de libérer des ressources pour améliorer d'autres aspects du processus, alors il y a là une réelle valeur ajoutée.

La dernière question est celle que j'ai vue revenir sans cesse dans le domaine de la santé publique : qu'en est-il de la longévité?

Nous avons observé ce phénomène dans plusieurs des projets et initiatives de notre plateforme *AI4PH*. Nous avons travaillé avec des étudiants et des étudiantes, des érudits et des stagiaires talentueux qui ont développé des applications innovantes et bien conçues. Mais que se passe-t-il une fois le projet terminé? Comment ce travail est-il maintenu, soutenu et intégré aux opérations courantes?

La surveillance des eaux usées en est un bon exemple. Certaines des premières innovations ont vu le jour dans des milieux universitaires, mais il faut finalement qu'il y ait une transition entre un projet de recherche fructueux et un programme de santé publique durable. Nous devons réfléchir à cette transition dès le début, et non pas après coup.

Sinon, nous risquons de mettre au point des solutions qui fonctionnent bien dans le cadre d'un projet pilote ou de recherche, mais qui ne peuvent être maintenues, transposées à plus grande échelle ou adoptées dans la pratique. En fin de compte, c'est un défi que la santé publique ne peut pas se permettre de relever.

Ainsi, si l'on considère certaines sources de données émergentes, j'ai inclus deux exemples ici.

Le premier concerne les données sur les eaux usées, que bon nombre d'entre nous ont appris à connaître pendant la pandémie de COVID. Ces données présentaient de nombreux aspects positifs. Elles étaient disponibles rapidement, couvraient une large partie de la population (du moins pour les collectivités raccordées à un réseau d'égouts) et soulevaient moins de préoccupations sur le plan de la protection de la vie privée et des aspects juridiques. Pas aucune, certes, mais certainement moins, car elles fournissaient une

estimation de l'infection à l'échelle de la population plutôt que des renseignements sur des personnes en particulier.

Cette approche s'est également révélée très exploitable. Nous avons vu ces indicateurs mis à profit pendant la pandémie de COVID, et, depuis, nous continuons à constater la valeur des données sur les eaux usées dans le cadre d'efforts plus larges de surveillance des maladies respiratoires. Dans de nombreux endroits, cela fait désormais partie des activités de surveillance de routine.

Cependant, il reste encore des questions importantes à se poser. Ce processus est-il reproductible? La collecte, l'analyse et l'interprétation des données peuvent-elles être menées de manière constante au fil du temps? Ce processus peut-il être durable et intégré aux opérations régulières de santé publique? Ces questions sont tout aussi importantes que l'innovation elle-même. Où ce processus est-il mis en œuvre? Qui valide continuellement ces modèles pour s'assurer qu'ils fonctionnent comme prévu au fil du temps?

En matière de surveillance des eaux usées, la réponse peut en réalité s'avérer assez complexe. La manière dont les échantillons sont prélevés a son importance. Différentes méthodes d'échantillonnage peuvent influencer les résultats, et chaque réseau d'égouts présente des caractéristiques propres dont il faut tenir compte dans les modèles. Il y a donc d'importantes questions méthodologiques à prendre en considération.

Parallèlement, les données sur les eaux usées constituent un bon exemple de situation où nous générons des renseignements véritablement nouveaux, des renseignements susceptibles d'améliorer la prise de décisions en matière de santé publique.

Si l'on se penche sur les données issues des médias sociaux, le tableau se complique quelque peu.

Nous savons qu'elles sont actualisées, ce qui est précieux. Elles sont aussi souvent plus rentables que de tenter de recueillir des renseignements sur l'opinion publique en interrogeant les gens un par un. De plus, elles peuvent fournir de nouvelles informations et perspectives auxquelles nous n'aurions peut-être pas accès autrement.

Cependant, cette approche comporte également des limites importantes. L'une des plus importantes est le biais de sélection. Tout le monde n'utilise pas les médias sociaux, et les personnes qui les utilisent ne sont pas nécessairement représentatives de l'ensemble de la population. Même parmi les utilisateurs et utilisatrices, certains groupes sont beaucoup plus enclins à publier, à interagir ou à donner leur opinion que d'autres.

Ainsi, bien que les données des médias sociaux puissent fournir des indications utiles, nous devons faire preuve d'une grande prudence pour bien comprendre qui est représenté dans ces données, qui ne l'est pas, et comment ces lacunes pourraient influencer les conclusions que nous tirons et les décisions que nous prenons.

Des questions persistent également quant à savoir s'il est approprié d'utiliser ces données en premier lieu et dans quelle mesure nous pouvons réellement tirer parti des renseignements qu'elles fournissent.

Nous pouvons probablement identifier des tendances émergentes ou des changements d'opinion, mais la question qui se pose alors est la suivante : quelles mesures devons-nous prendre? Comment ces renseignements se transformant-ils en une intervention de santé publique?

Cela ne veut pas dire qu'ils ne peuvent pas être utiles, au contraire. Cependant, je soulève ce point pour rappeler que chaque fois que nous envisageons une nouvelle source de données, nous devons réfléchir de manière critique à la valeur qu'elle apporte et à notre capacité à la maintenir et à l'utiliser efficacement à long terme.

Plus nous ferons preuve de rigueur en nous posant ces questions, plus nous aurons de chances de concentrer nos efforts sur les sources de données et les applications qui créent véritablement de la valeur. Et je pense que c'est là que nous commencerons à voir les avantages se concrétiser beaucoup plus rapidement. À l'heure actuelle, une partie de cette valeur est peut-être diluée, car nous explorons trop de possibilités à la fois. En nous concentrant sur les approches qui génèrent des informations pertinentes et qui favorisent de meilleures décisions, nous serons en mesure de concrétiser cette valeur de façon beaucoup plus claire.

[Cette diapositive illustre comment des biais peuvent s'introduire tout au long du cycle de vie de l'IA, depuis la sélection du problème et la collecte de données jusqu'au déploiement et à l'utilisation.] Je pourrais passer des heures à parler de l'un des plus grands défis liés à l'IA : les biais algorithmiques, et plus précisément, la manière dont ils peuvent nuire à l'équité.

Comme je l'ai mentionné au début, c'est un sujet qui reviendra tout au long de la série de conférences, car il est pertinent dans presque tous les domaines où l'IA est appliquée. Alors, par souci de concision, je me concentrerai sur quelques points clés.

Le premier est que les biais présents dans les données peuvent se propager par l'entremise des algorithmes, et que ces algorithmes peuvent ensuite influencer les décisions de manière à renforcer, voire à amplifier, les iniquités existantes. C'est une préoccupation majeure.

Il est également important de reconnaître qu'il ne s'agit pas d'un défi qui commence et se termine avec l'IA. Des biais peuvent émerger chaque fois que nous recueillons, analysons ou utilisons des données pour orienter nos décisions. La santé publique est aux prises avec ces enjeux depuis longtemps, et nous sommes très conscients des risques.

Ce qui a changé, c'est que le débat s'éloigne de plus en plus de la simple identification des biais pour se concentrer sur la manière de les atténuer. Quelles mesures de protection peut-on mettre en place? Comment détecter les biais tôt? Comment évaluer si un système fonctionne de manière équitable pour différentes populations? Et que faire lorsque ce n'est pas le cas?

Ce sont des questions qui animent désormais une grande partie du débat, et elles sont essentielles si nous voulons mettre en œuvre l'IA d'une manière efficace et équitable.

Si l'on repense à 2019 et au début des années 2020, le débat portait en grande partie sur la sensibilisation. Nous tirions la sonnette d'alarme pour souligner que les biais étaient importants, qu'ils pouvaient influencer les résultats et qu'il fallait en tenir compte à chaque utilisation de l'IA. Et cela est encore important.

Le débat a évolué. Aujourd'hui, l'accent est beaucoup plus sur les solutions. Quelles méthodes pouvons-nous utiliser pour repérer et atténuer les biais? Comment intégrer ces considérations dans l'ensemble du processus plutôt que de les traiter comme une réflexion après coup?

C'est pour cette raison que j'aime réfléchir à l'équité tout au long du processus d'IA. Cela commence par une question très fondamentale : quels problèmes choisissons-nous de résoudre, et pourquoi?

À bien des égards, c'est là que commence l'équité. Avant de parler de données, de modèles ou d'algorithmes, nous devons nous demander quels besoins sont prioritaires et lesquels risquent d'être négligés.

Lorsque nous recueillons des données ou que nous utilisons des données existantes, nous devons nous demander où des biais pourraient se cacher dans ces données. Comment définissons-nous les résultats que nous essayons de prédire? Quels résultats privilégions-nous? Et comment élaborons-nous le modèle lui-même?

En fait, il existe des pratiques, au cours du développement d'un modèle, qui peuvent accroître le risque de biais si elles ne sont pas prises en compte. D'un autre côté, il existe également des approches bien établies pour aider à atténuer ces risques, comme l'examen de la performance au sein de différents sous-groupes et la validation sur plusieurs sites et auprès de plusieurs populations.

Je n'ai pas le temps d'aborder toutes ces méthodes aujourd'hui, mais il existe un certain nombre de pratiques exemplaires reconnues qui ont émergé dans ce domaine.

Et même si tout se passe bien, des biais peuvent encore s'introduire une fois que le modèle est mis en pratique. La façon dont un modèle est utilisé peut être tout aussi importante que la façon dont il a été élaboré.

Imaginons que nous disposions d'un modèle très performant. Il est efficace et fonctionne de manière constante dans différents sous-groupes pour identifier les personnes susceptibles de bénéficier d'une intervention particulière. À bien des égards, c'est le scénario idéal.

Mais même dans ce cas, des biais peuvent apparaître dans la façon dont le modèle est utilisé. S'il n'est accessible qu'à un sous-ensemble de la population, ou s'il n'est utilisé que par des personnes le connaissant, y ayant accès ou ayant les moyens de se le permettre, on continuera d'observer des résultats inéquitables.

Dans ce cas, le biais ne provient pas nécessairement du modèle lui-même. Il découle de la façon dont le modèle est mis en œuvre et dont ses avantages sont répartis.

Il s'agit donc d'un sujet complexe. Mais ce que je tiens à souligner, c'est que la santé publique a un rôle crucial à jouer pour relever ces défis, notamment en reconnaissant

quand un modèle ne devrait pas être utilisé du tout. Nous disposons d'une grande expertise à apporter à ces questions, et cette expertise est nécessaire à chaque étape, du développement à la mise en œuvre.

Ce qui m'amène au troisième domaine que je souhaitais aborder à ce sujet, soit l'empreinte environnementale de l'IA.

Il s'agit d'un enjeu complexe, et c'est un domaine où, en santé publique, nous avons tendance à réfléchir attentivement aux compromis qu'il implique.

Nous savons que l'IA a une empreinte environnementale, et celle-ci ne se résume pas à un seul aspect. Elle commence par le matériel lui-même, notamment l'extraction des matières premières et des minéraux rares, ainsi que la fabrication nécessaire à la mise en place de l'infrastructure informatique. Viennent ensuite l'énergie requise pour l'entraînement et l'exécution des modèles, les UTG et autres infrastructures informatiques spécialisées qui les alimentent, ainsi que le réseau croissant de centres de données qui soutiennent l'ensemble de ce système.

Les centres de données nécessitent évidemment une infrastructure importante et doivent être refroidis, ce qui entraîne des besoins supplémentaires en énergie et en ressources. Et à mesure que l'adoption de l'IA continue de croître, tous ces facteurs doivent être pris en compte dans le débat.

Ainsi, lorsque nous réfléchissons à l'IA, il est important de considérer non seulement les avantages potentiels, mais aussi les coûts environnementaux associés au développement et à l'exploitation de ces systèmes.

Et plus nous utilisons ces technologies, plus leur impact environnemental est important, qu'il s'agisse de la consommation d'électricité, de l'utilisation de l'eau ou d'une série d'effets locaux.

Nous sommes également meilleurs à mesurer ces impacts. Par exemple, des travaux récents ont examiné comment les infrastructures liées à l'IA contribuent aux émissions de carbone, et il existe désormais des estimations de l'empreinte carbone pour de nombreuses technologies d'IA. Notre compréhension devient donc beaucoup plus approfondie.

Et puis, même après tout cela, il faut tenir compte du cycle de vie complet de la technologie. Le matériel doit éventuellement être remplacé, mis à niveau et éliminé, ce qui engendre des répercussions environnementales supplémentaires qui vont bien au-delà de l'utilisation des modèles eux-mêmes.

Évidemment, il y a aussi les déchets électroniques et tous les défis qui en découlent.

Cela crée une véritable tension pour la santé publique. D'une part, certains de ces outils peuvent nous aider à relever les défis liés au climat. Ils peuvent faciliter la détection précoce des feux de forêt, aider à identifier les populations à risque de déplacement lié au climat et générer des gains d'efficacité susceptibles de réduire les émissions ou la consommation de ressources dans d'autres parties du système.

L'utilisation de ces technologies, lorsqu'elle est bien menée, présente donc des avantages environnementaux potentiels. Parallèlement, leur développement, leur déploiement et leur exploitation entraînent des coûts environnementaux.

Le défi consiste à trouver le juste équilibre. Nous devons comprendre dans quels domaines l'IA crée une valeur réelle et dans quels cas cette valeur est suffisamment importante pour justifier son empreinte environnementale. Comme pour bon nombre des enjeux dont nous avons discuté aujourd'hui, il ne s'agit pas simplement d'une question d'avantages ou d'inconvénients. Il s'agit de peser soigneusement les deux aspects et de prendre des décisions éclairées quant à la manière et aux endroits où ces technologies devraient être utilisées.

[Cette diapositive présente un diagramme illustrant les stratégies visant à réduire les répercussions environnementales de l'IA, notamment l'apprentissage par transfert, l'élagage et la quantification ainsi que l'utilisation de modèles plus petits et spécifiques à une tâche.] Alors, que faire face à cela? C'est un véritable casse-tête pour nous.

J'aborde la question sous trois angles.

D'abord, nous devons reconnaître qu'il existe des mesures que nous pouvons prendre pour réduire l'empreinte environnementale de l'IA. Dans certains cas, cela peut signifier ne pas utiliser du tout l'IA. C'est pourquoi je reviens sans cesse sur l'importance d'identifier les applications qui créent le plus de valeur pour la santé publique et de concentrer nos efforts sur celles-ci, plutôt que de considérer l'IA comme une technologie qui devrait être appliquée partout.

Il existe également des approches qui peuvent s'avérer utiles. Je ne vais pas entrer dans les détails, mais l'apprentissage par transfert en est un exemple. Il s'agit d'un concept très élégant qui consiste à prendre un modèle déjà entraîné pour une tâche donnée et à l'adapter à une autre.

Pourquoi est-ce important? Parce que cela signifie que nous n'avons pas à repartir de zéro chaque fois. L'entraînement d'un nouveau modèle peut demander beaucoup de ressources. Par conséquent, la réutilisation et l'adaptation de modèles existants peuvent réduire considérablement les répercussions environnementales. Nous l'avons constaté à maintes reprises, et dans bien des cas, cela devrait tout simplement être considéré comme une pratique exemplaire.

Il existe d'autres approches également. Des techniques comme l'élagage et la quantification reposent sur l'idée qu'il n'est pas toujours nécessaire d'utiliser toutes les variables ou tous les paramètres d'un modèle pour obtenir de bonnes performances. Cela peut sembler contre-intuitif, mais en rendant les modèles plus efficaces, nous pouvons souvent réduire leur empreinte environnementale.

On observe également un intérêt croissant pour l'utilisation de modèles plus petits, adaptés à des tâches spécifiques, plutôt que de se fier à de grands modèles polyvalents pour tout. Dans certains cas, un modèle plus petit peut accomplir la tâche tout aussi efficacement tout en nécessitant beaucoup moins de ressources.

Je ne dis pas que toutes ces approches fonctionneront dans toutes les situations, mais elles démontrent qu'il existe des moyens concrets de réduire les répercussions environnementales de l'IA.

Le même principe s'applique à la façon dont nous concevons nos flux de travail. Nous n'avons pas toujours besoin de réentraîner un modèle à partir de zéro. Parfois, nous pouvons mettre à jour, adapter ou réutiliser des modèles existants de manière à obtenir le même résultat avec beaucoup moins de ressources.

Il y a donc des occasions d'adopter une approche beaucoup plus réfléchie quant à la manière dont nous développons et déployons ces technologies, et cela devrait faire partie de la discussion dès le départ.

Nous devrions également mettre en place des processus d'évaluation rigoureux afin de comprendre quels modèles fonctionnent, dans quels contextes et pour qui.

De nombreux efforts sont également déployés pour rendre les centres de données et l'infrastructure informatique plus efficaces. Et surtout, nous devons éviter toute utilisation inutile, comme je l'ai mentionné plus tôt.

Le dernier point concerne la gouvernance.

Lorsque nous parlons de gouvernance de l'IA, nous commençons souvent par les données et la confidentialité. Nous pensons à l'éthique, aux biais et à l'équité. Tous ces aspects sont importants, mais nous devons également commencer à réfléchir aux répercussions environnementales.

En d'autres mots, la durabilité environnementale doit également faire partie du débat sur la gouvernance. Elle devrait constituer l'un des facteurs à prendre en compte lorsque nous décidons comment et où utiliser l'IA.

Comme nous le faisons dans de nombreux autres aspects de la santé publique, nous devons commencer à nous demander s'il s'agit de la solution la plus durable sur le plan environnemental, ou s'il en existe de meilleures.

Il existe des moyens d'évaluer cela. Dans le cours que je mentionnerai à la fin, nous passons en revue certaines de ces approches et certains de ces calculs. Je mentionne cela simplement pour illustrer qu'il ne s'agit pas d'une discussion purement théorique. Il existe des moyens concrets d'évaluer ces compromis.

Passons donc à la dernière partie de la présentation : de quoi avons-nous réellement besoin en santé publique?

J'ai parlé des capacités de l'IA, présenté quelques exemples et mis en évidence certains domaines où nous devons réfléchir de manière critique à ses répercussions. La question qui se pose maintenant est la suivante : à quoi ressemble réellement un déploiement réussi?

Cela nous ramène à un point que j'ai délibérément répété à plusieurs reprises. Avant tout, l'IA doit répondre aux objectifs de santé publique, mais, au-delà de cela, nous devons également nous assurer qu'elle est efficace, sécuritaire et éthique.

[Cette diapositive présente un cadre décrivant les conditions pour une mise en œuvre réussie de l'IA : harmonisation avec les objectifs de santé publique, efficacité, sécurité, éthique, équité, confidentialité et durabilité.]

J'aimerais aborder brièvement la manière dont nous pouvons garantir que cela se produise, en précisant toutefois que cette responsabilité n'incombe à personne en particulier. Des structures organisationnelles et des processus doivent être mis en place pour soutenir cette démarche.

En 2022, nous avons publié un article examinant les domaines prioritaires sur lesquels les organisations de santé publique doivent se concentrer pour mettre l'IA en pratique. Vous y verrez plusieurs thèmes que j'ai déjà abordés, notamment la gouvernance, l'infrastructure, et l'évaluation. Deux autres thèmes qui me semblent particulièrement importants sont le développement et la formation de la main-d'œuvre, ainsi que les partenariats. Si nous avons du temps à la fin, j'y reviendrai.

Ma réflexion a également un peu évolué depuis. Lorsque les gens parlent d'infrastructure, ils pensent souvent à celle qui est nécessaire à l'exécution des modèles, mais nous avons également besoin d'une infrastructure pour les essais, la surveillance et l'évaluation continue.

Il y a quelque chose après le déploiement. Les modèles doivent être surveillés, mis à jour et évalués au fil du temps. Bien qu'il existe désormais de nombreuses normes et pratiques exemplaires en matière d'IA, nous devons également nous assurer que les gens disposent du temps et des capacités nécessaires pour les appliquer. Cet examen critique ne se fait pas automatiquement. Il doit être intégré au processus et, éventuellement, officialisé.

Je pense donc que le cadre tient toujours la route, avec quelques ajouts et améliorations fondés sur ce que nous avons appris au cours des dernières années.

[Cette diapositive présente un diagramme illustrant la mise en œuvre responsable de l'IA, soutenue par l'engagement, la supervision, la réglementation, les directives et les normes.] Au cœur de tout cela, je vois deux piliers essentiels.

D'une part, il y a la réglementation, les directives et les normes, qui évoluent chaque jour. Elles sont importantes et nous en avons besoin.

Cependant, nous ne pouvons pas nous fier uniquement à elles. Au Canada, nous avons un peu de retard dans certains domaines, même si de nombreuses personnes y travaillent; je suis donc convaincue que nous rattraperons notre retard.

D'autre part, il y a l'engagement : l'engagement auprès des personnes touchées par ces technologies, des praticiens et des praticiennes qui les utilisent, ainsi que des membres de la direction et des responsables du système de santé chargés de leur mise en œuvre. Cet engagement doit être important et authentique.

À bien des égards, c'est déjà ainsi que nous fonctionnons en santé publique. Nous avons de solides traditions d'approches communautaires et participatives, et ces mêmes principes devraient guider notre façon de concevoir, de développer et de déployer l'IA.

Il y a donc plusieurs couches. Il y a des individus, comme tous ceux qui participent à cet appel, qui en apprennent davantage sur ces technologies et leur fonctionnement. Il y a des organisations qui mettent en place des pratiques, des politiques et des structures de gouvernance, puis il y a des règlements et des normes de niveau supérieur.

La santé publique adore le modèle du fromage suisse, et c'est exactement ainsi que je vois les choses. Nous avons besoin de multiples niveaux de protection. Les individus ont des responsabilités. Les organisations ont des responsabilités. Nous avons besoin de politiques, de normes, de suivi et de supervision.

Aucune couche à elle seule, qu'il s'agisse de la réglementation ou quoi que ce soit d'autre, ne permettra d'éliminer tous les risques dont nous avons parlé aujourd'hui. C'est la combinaison de ces différentes couches, ensemble, qui permet de les gérer.

Par souci de concision, je ne m'attarderai pas sur certains cadres précis. Cependant, je tiens à souligner qu'un travail considérable a été accompli ces dernières années en matière d'éthique, d'équité et d'IA. Il existe de nombreux cadres excellents, ce qui signifie que nous ne partons pas de zéro. Nous disposons déjà d'une base solide sur laquelle nous appuyer.

[Exemple de projet participatif d'IA pour la prévention du diabète illustrant la collaboration entre chercheurs et chercheuses, praticiens et praticiennes, et partenaires communautaires.]

Et pour terminer par un exemple concret, je vais vous parler d'un projet sur lequel nous travaillons, qui porte sur l'utilisation de l'IA pour la prédiction du risque de diabète.

Ce que je tiens à souligner, c'est que le cœur de ce projet ne réside pas dans la technologie en soi. La technologie est là. Le modèle existe, et nous continuons à l'améliorer, mais c'est presque secondaire.

Ce sur quoi nous nous sommes vraiment concentrés, c'est l'engagement ciblé : comment le modèle est utilisé, comment les résultats sont présentés et comment ils orientent les politiques et la prise de décision. Et cela a été une expérience d'apprentissage extraordinaire.

Si la façon dont ce processus d'engagement a façonné le projet et modifié certaines de nos approches vous intéresse, plusieurs articles sont disponibles.

Ce qui s'est avéré particulièrement précieux, c'est que bon nombre des considérations relatives à une IA responsable ont été mises en lumière grâce à ces conversations. Les questions concernant la mise en œuvre, l'équité, la transparence et les répercussions émergent toutes naturellement lorsque l'on parle avec les personnes qui, au bout du compte, utiliseront cette technologie ou en subiront les effets.

[Cette diapositive présente un diagramme du cycle de vie illustrant les étapes d'un déploiement responsable de l'IA : définir, développer, préparer, évaluer, suivre et adapter.] Et je terminerai par une dernière réflexion sur le cycle de vie du déploiement.

Je ne voudrais pas que quiconque quitte cette conférence en pensant qu'il suffit de choisir un outil d'IA, de le mettre en pratique, et que c'est tout. Ça ne fonctionne pas ainsi.

Il s'agit d'un cycle. Une fois qu'un système d'IA est déployé, même après tout le travail minutieux de conception, de codéveloppement et de mobilisation, cet engagement doit se poursuivre. Nous devons continuer à tirer des leçons des personnes qui utilisent la technologie et de celles qui en subissent les effets.

Nous devons également faire preuve de beaucoup plus de rigueur dans l'évaluation des résultats, idéalement de manière prospective. Cela commence à se produire, mais les progrès ont été relativement lents.

Ensuite, les enseignements tirés du suivi et de l'évaluation doivent être réinjectés dans le processus, que ce soit pour améliorer l'application actuelle ou pour nous aider à résoudre plus efficacement le prochain problème.

Le déploiement ne marque donc pas la fin du processus. À bien des égards, il s'agit du début de la prochaine étape.

Voici ma dernière diapositive, et je terminerai par un point qui me semble particulièrement important pour la santé publique.

Selon moi, l'évaluation est la prochaine priorité. La question n'est pas simplement de savoir si nous adoptons l'IA. La question est de savoir comment nous démontrons que l'IA apporte réellement une valeur ajoutée.

Pour y répondre, nous devons regarder au-delà de la performance des modèles. Évidemment, nous devons savoir si la technologie fonctionne, mais nous devons également comprendre comment elle modifie les pratiques, quels avantages elle procure en matière de santé publique, et quelles en sont les répercussions sur l'équité et la durabilité, sujets que j'ai abordés tout au long de cette présentation.

Je terminerai donc en disant que nous avons abordé de nombreux sujets aujourd'hui. Nous avons parlé des fondements de l'IA, de certaines de ses applications en santé publique, des répercussions dont nous devons tenir compte et, brièvement, de ce à quoi pourrait ressembler une mise en œuvre responsable.

La bonne nouvelle, c'est que tous les sujets que j'ai abordés aujourd'hui seront approfondis tout au long du reste de la série de conférences. J'ai vu ce qui nous attend, et bon nombre de ces sujets seront examinés de manière beaucoup plus approfondie.

J'ai pensé terminer en mentionnant brièvement ce cours. L'an dernier, à la *Dalla Lana School of Public Health*, nous avons offert notre premier cours sur l'introduction à l'IA en santé publique. Nous avons consacré beaucoup plus de temps à explorer bon nombre des sujets dont nous avons discuté aujourd'hui. C'était un groupe de participants et de participantes incroyablement réfléchis, issus d'un large éventail de disciplines.

À la fin du cours, je leur ai posé quelques questions : quelles possibilités vous emballent le plus? Qu'est-ce qui vous préoccupe? Et que devons-nous faire pour nous assurer que tout se passe bien pour la santé publique?

Ensemble, les participants et les participantes ont cerné un ensemble de thèmes, et nous les avons transposés sur ce signet en guise de rappel des conversations que nous avons eues.

Je pense que je vais m'arrêter ici. Il se passe beaucoup de choses dans le domaine de l'IA, mais l'avenir n'est pas encore tracé. Nous avons l'occasion de façonner la manière dont ces technologies seront utilisées en santé publique.

Pour ceux et celles qui travaillent dans le domaine depuis plusieurs années, je vous invite à vous joindre à nous pour aider à définir ce qu'est une utilisation responsable. Nous ne devrions pas partir du principe que quelqu'un d'autre prendra ces décisions à notre place.

Comme je l'ai dit à plusieurs reprises aujourd'hui, je crois que la santé publique a un rôle unique et important à jouer pour veiller à ce que ces technologies soient utilisées de manière à faire progresser les objectifs de santé publique et à générer des résultats positifs.

Je m'arrête donc ici. Je suis certaine que vous avez beaucoup de questions et de réflexions, et j'aimerais beaucoup les entendre. Merci beaucoup de m'avoir invitée.

Marc-André Bélair :

Super! Merci beaucoup, Dre Rosella, pour votre présentation captivante et équilibrée. Une question que j'aimerais poser, c'est que vous avez insisté sur le fait de partir d'un problème de santé publique, mais comment une organisation peut-elle déterminer quand il vaut la peine de s'orienter vers l'IA et quand une approche plus simple serait plus adaptée à ces besoins?

Dre Laura Rosella :

C'est une très bonne question. En passant, je dirais que c'est là-dessus que nous devrions consacrer le plus de temps, car il s'agit sans doute de la décision la plus importante de tout le processus.

En pratique, les organisations abordent cette question de différentes façons. Certains services de santé et certaines organisations passent beaucoup de temps à définir le problème au préalable. D'autres fois, une technologie fait son apparition et quelqu'un dit « Je pense que cela pourrait vraiment nous aider », puis nous essayons de l'intégrer à un système existant.

Je commencerais par cerner les décisions que vous souhaitez améliorer ou les limites de l'approche actuelle. Qu'est-ce qui ne fonctionne pas aussi bien que prévu? Le processus est-il trop lent? Certains groupes sont-ils laissés pour compte? Y a-t-il des lacunes dans les informations sur lesquelles nous nous appuyons pour prendre des décisions? Mettez d'abord ces problèmes en évidence.

Comparez ensuite les solutions possibles. De quelles différentes façons pourrions-nous améliorer la situation?

Parfois, en effectuant cet exercice, on se rend compte qu'il existe des solutions qui n'impliquent pas du tout l'IA. Il existe parfois des méthodes plus simples qui fonctionneraient tout aussi bien. Cependant, s'il est clair que l'IA pourrait améliorer le processus de manière importante, alors c'est un excellent point de départ.

L'exemple des tours de refroidissement est un bon exemple. Nous avions un problème clair : l'approche existante était lente, nous savions que certains emplacements nous échappaient, et il arrivait que certains éléments apparaissent sur la carte une année pour disparaître la suivante. Le cas d'utilisation était très clair, tout comme la décision qui s'en dégagerait pour améliorer la situation.

Je pense également qu'il est important de déterminer ce qui va donner suite à ces renseignements. Quel est le flux de travail complet? La composante d'IA n'en représente généralement qu'une petite partie. Le plus grand défi consiste à comprendre comment ces renseignements circulent dans le système et comment ils contribuent, en fin de compte, à une décision ou à une mesure.

Cela me rappelle d'ailleurs un atelier animé en 2022 par l'un de nos mentors du programme *AI4PH*, le professeur Dan Lizotte. Dan est un informaticien, mais il a réuni des professionnels locaux en matière de santé publique et des spécialistes de premier plan en IA, et il a commencé par poser une question simple : « Quels sont vos problèmes? ».

L'objectif n'était pas de partir la technologie. Il s'agissait d'abord d'écouter. Ensuite, les spécialistes en IA pouvaient répondre en disant : « Voici une approche que nous avons déjà utilisée et qui pourrait vous aider » ou « Voici une méthode qui pourrait être pertinente dans ce contexte ».

Je trouve que c'est un excellent modèle. Partir des défis locaux en matière de santé publique, puis réunir des gens pour réfléchir ensemble à des solutions possibles est souvent l'un des meilleurs moyens de déterminer où l'IA peut véritablement apporter une valeur ajoutée.

Marc-André Bélair :

Super, merci beaucoup. Une question qui nous vient de Slido est : quels modèles d'IA utilisez-vous actuellement dans le cadre de votre travail et où sont-ils hébergés?

Sont-ils hébergés localement ou au sein d'une organisation privée?

Dre Laura Rosella :

Oui, bonne question.

Dans le cadre de mon travail, je m'occupe beaucoup de prévision et de surveillance. Dans l'ensemble, nous utilisons des approches d'apprentissage automatique plus traditionnelles, comme des modèles d'ensemble et des arbres améliorés.

Dans bien des cas, ces modèles sont hébergés localement, car ils sont construits à partir de données qui se trouvent dans un environnement protégé. Cela a été un facteur important pour nous, et je sais que c'est un défi auquel de nombreuses organisations sont également confrontées.

En général, les modèles sont déployés au sein d'une infrastructure organisationnelle sécurisée. Il s'agit parfois d'une infrastructure locale, et parfois d'une infrastructure sécurisée dans le nuage. Lorsque des environnements dans le nuage sont utilisés, ils sont

généralement gérés par des organisations qui disposent de l'autorité et des mesures de protection nécessaires pour traiter de manière appropriée les renseignements personnels sur la santé.

C'est toutefois une bonne question, car ce contexte évolue rapidement.

Plus généralement, une grande partie du débat actuel sur l'IA au Canada porte sur la souveraineté. Comment s'assurer que l'infrastructure qui soutient ces technologies est sécurisée, protégée et conforme aux exigences canadiennes? Je sais que c'est une question à laquelle de nombreuses organisations réfléchissent activement lorsqu'elles prennent des décisions concernant l'adoption et le déploiement de l'IA.

Marc-André Bélair :

Super. Merci beaucoup. Maintenant, je vois une question ici et je constate que de nombreuses équipes au sein du gouvernement du Canada considèrent l'avènement de l'IA agentive comme le prochain grand tournant. Selon vous, que doivent faire les fonctionnaires, et peut-être aussi les professionnels et professionnelles de la santé publique qui sont fonctionnaires, pour se préparer à un monde où l'IA agentive sera de plus en plus intégrée aux processus de travail?

Dre Laura Rosella :

Oui, j'adore cette question. Merci de l'avoir posée.

Pour ceux et celles qui entendent ce terme pour la première fois, rappelons que l'IA agentive désigne des systèmes qui ne se contentent pas d'exécuter une seule tâche. Ils sont capables de coordonner une série d'actions pour atteindre un objectif. Ainsi, au lieu de simplement identifier une personne à risque, par exemple, un agent pourrait identifier cette personne, prendre un rendez-vous, générer les documents nécessaires et effectuer plusieurs étapes du processus.

Il y a quelques points importants à prendre en compte.

D'abord, il y a des répercussions évidentes sur la main-d'œuvre. À mesure que les agents d'IA deviendront plus aptes à accomplir des tâches actuellement effectuées par des humains, d'importantes questions se poseront quant à leurs implications pour les différents rôles et professions, et cela ne se limite pas à la santé publique. On observe déjà des exemples d'agents de recherche capables d'aider à formuler des questions de recherche, à mener des analyses et à rédiger des articles. Ces évolutions se produisent dans de nombreux secteurs.

C'est pour cette raison que je pense qu'il est important de prêter attention à ces tendances et de comprendre où elles mènent. C'est également l'occasion de réfléchir aux domaines où les humains apportent le plus de valeur ajoutée et où leur implication demeure essentielle.

Je pense également qu'il est important de se demander dans quels domaines le public s'attend à ce que les humains continuent de jouer un rôle et souhaite qu'ils le fassent. C'est une discussion que nous devons mener parallèlement à celle sur la technologie.

Le deuxième point est que, bien que l'IA agentive suscite beaucoup d'enthousiasme, les données d'évaluation sont encore relativement limitées. Je participe actuellement à des travaux d'évaluation de systèmes agentiques, et, lorsque l'on examine la documentation, on constate qu'elle est encore assez limitée.

Ainsi, bien que ces technologies soient manifestement importantes et méritent toute notre attention, je ne pense pas que nous en soyons encore au point de pouvoir les intégrer en toute confiance dans des flux de travail complexes et de supposer qu'elles fonctionneront comme prévu. Nous avons besoin de beaucoup plus de recherche et d'évaluation, en particulier en ce qui concerne les processus de bout en bout.

C'est une chose de démontrer qu'un modèle donne de bons résultats pour une tâche précise, mais ça en est une autre de comprendre les répercussions lorsque ce modèle s'intègre à un processus en plusieurs étapes qui influence les décisions et les actions.

Je conseillerais donc de rester à l'affût de ce domaine et de prêter attention aux données qui émergent, en particulier aux évaluations rigoureuses des systèmes agents. Une grande partie de ce travail se déroule actuellement dans le secteur des soins de santé, et dans une moindre mesure en santé publique.

Parallèlement, il faut rester concentré sur les aspects humains du travail. Nous devons déterminer clairement dans quels domaines la supervision, les relations et le jugement humain sont importants, et veiller à ce que ces éléments ne soient pas perdus de vue à mesure que ces technologies évoluent.

Marc-André Bélair :

Oui, c'est un aspect vraiment important à prendre en compte.

J'aimerais changer légèrement de sujet, car je pense que cela rejoint un point que vous avez abordé plus tôt, ainsi qu'un point que j'ai soulevé dans ma déclaration de reconnaissance territoriale.

L'une des préoccupations que l'on entend concernant l'IA est que les données d'entraînement sur lesquelles reposent de nombreux modèles peuvent être culturellement homogènes. Cela soulève une question importante pour des organisations comme la nôtre : dans quelle mesure ces modèles sont-ils représentatifs des communautés que nous servons réellement?

Et je ne parle pas seulement de la population canadienne en général. Je pense plus précisément aux communautés autochtones, aux communautés non anglophones et aux autres populations dont les expériences ne sont peut-être pas bien représentées dans les données utilisées pour entraîner ces systèmes.

Alors, comment devrions-nous aborder les biais qui peuvent exister dans les données d'entraînement et les ensembles de données? Et comment évaluons-nous si les systèmes d'IA représentent de manière appropriée les populations auprès desquelles nous les utilisons?

Quelqu'un dans le clavardage a également mentionné l'importance de s'assurer que l'IA soit culturellement intelligente, à l'image de certains travaux menés par Tulong Tech.

J'aimerais connaître votre point de vue à ce sujet. Comment pouvons-nous concevoir des systèmes d'IA à la fois représentatifs et adaptés aux diverses communautés que nous servons?

Dre Laura Rosella :

Je pense que c'est une question très importante.

Je vais souligner quelques points.

D'abord, des érudits, des communautés et des leaders autochtones ont déjà mené de nombreux travaux approfondis dans ce domaine. Par exemple, il existe des articles de position et des documents d'orientation sur les approches autochtones en matière d'intelligence artificielle, et je pense qu'il est important que nous reconnaissions ces travaux, que nous en tirions des leçons et que nous veillions à ce qu'ils se reflètent dans notre approche de la gouvernance et de la mise en œuvre.

Plus généralement, nous disposons déjà de cadres de référence, tels que les principes PCAP, les principes CARE et d'autres approches qui nous aident à réfléchir à la gouvernance et à la gestion responsable des données. Cela ne veut pas dire que le travail est terminé, mais nous comprenons déjà bon nombre des défis liés aux données, avec ou sans IA.

La première chose que je tiens donc à souligner est que nous devons mettre ces principes au premier plan. Quoi que nous fassions avec l'IA, nos actions devraient s'inspirer de ces principes dès le départ.

Le deuxième point est que nous devons entamer ces conversations beaucoup plus tôt. Avant d'aborder les questions relatives aux algorithmes, aux biais et à la performance des modèles, nous devrions nous poser la question suivante : quel problème essayons-nous de résoudre? Pourquoi choisissons-nous ce problème? Et quelles répercussions cela pourrait-il avoir sur les communautés concernées?

Je pense que c'est là que doivent avoir lieu certaines des discussions les plus importantes.

Nous devons également mieux comprendre comment les modes de savoir autochtones peuvent façonner ces discussions. Il existe déjà des idées et des cadres solides. Là où je vois encore une marge de progression possible, c'est dans la cohérence avec laquelle ils sont appliqués dans la pratique.

Pour moi, cela va au-delà de la gouvernance. Il s'agit de la manière dont nous abordons cette transformation plus large dès le départ.

En ce qui concerne votre question précise sur les biais dans les données et la représentation, je pense que ces aspects doivent de plus en plus faire partie intégrante du processus d'évaluation et d'adoption des systèmes d'IA.

Avant d'utiliser un outil d'IA, nous devrions nous poser des questions telles que : A-t-il été validé auprès des populations qui comptent pour notre communauté? Qui est représenté dans les données? Qui n'y est pas représenté? Et quelles sont les implications de ces lacunes?

Selon les réponses, nous pourrions décider que l'outil n'est pas approprié, ou qu'il ne peut être utilisé qu'avec des mises en garde et des mesures de protection importantes.

À bien des égards, ce sont des questions que nous avons déjà l'habitude de poser en santé publique. Nous devons simplement nous assurer que cette même rigueur s'applique également à l'IA.

La seule chose que j'ajouterais, c'est que je ne veux pas que cela devienne uniquement un débat sur la gouvernance. La gouvernance est importante, mais elle ne peut pas être la seule mesure de protection. Nous ne pouvons pas supposer que les processus de gouvernance permettront de détecter tous les problèmes si nous n'avons pas pris en compte ces questions lors de la définition du problème, de la conception de l'approche et de l'élaboration du modèle.

Cela doit vraiment être pris en compte tout au long du cycle de vie.

Marc-André Bélair :

Non, absolument. Merci. Et nous avons malheureusement bien plus de questions que nous pourrons en répondre aujourd'hui, ce qui, je tiens à le dire, me fait très plaisir, car cela montre que les gens sont intéressés. Mais si nous n'avons pas le temps de répondre à votre question aujourd'hui, je m'en excuse. Il y a toutefois ici une question que je trouve vraiment intéressante concernant l'IA et ce débat sur le remplacement du personnel de niveau d'entrée. Celle-ci porte plus précisément sur le fait de savoir si l'IA est censée remplacer les postes de niveau d'entrée comme ceux nécessitant une formation, comme une maîtrise en santé publique ou une maîtrise en épidémiologie publique, et comment cela va réorienter nos responsabilités vers la supervision, l'engagement communautaire et l'équité algorithmique? Quelles compétences nous permettront de rester indispensables?

Dre Laura Rosella :

Oui, bonne question. C'est tellement une bonne question, et honnêtement, j'y pense beaucoup.

Je ne suis sans doute pas tout à fait impartiale, puisque je travaille dans le domaine de la santé publique, mais je pense que ce secteur est confronté à cette tension plus que tout autre.

Tout d'abord, je tiens à préciser que, dans le cadre de certains de nos projets de recherche, nous avons effectivement testé cette hypothèse. Nous avons comparé des approches fondées sur l'IA, des approches humaines et des approches hybrides pour des tâches typiques de santé publique. D'après ces travaux, je ne pense pas qu'il s'agisse simplement de remplacer les gens.

Ce que je pense que l'on observe, c'est plutôt un changement dans la manière dont les gens répartissent leur temps. Peut-être consacrons-nous moins de temps à certaines taches techniques ou répétitives, mais davantage à la supervision, à la validation, à l'interprétation, au jugement et à la vérification que les résultats sont utilisés de façon appropriée.

C'est ainsi que je vois les choses. Il s'agit moins d'un remplacement que d'un changement dans la nature même du travail.

La raison pour laquelle la santé publique doit prendre cela au sérieux est que l'emploi en soi est un déterminant de la santé. Ainsi, lorsque nous nous demandons si l'IA pourrait remplacer les postes de niveau d'entrée, nous ne posons pas seulement une question relative à la main-d'œuvre. Nous posons également une question relative à la santé de la population.

Que se passerait-il si les gens avaient moins d'occasions d'acquérir de l'expérience au début de leur carrière? Quelles en seraient les répercussions sur le développement de la main-d'œuvre et, en fin de compte, sur la santé de la population? Ce sont là des questions sur lesquelles nous devons nous pencher sérieusement.

Pour ceux et celles qui sont actuellement en formation ou en début de carrière, je vous conseillerais de vous familiariser autant que possible avec ces technologies. Non pas parce que vous devrez devenir des spécialistes de l'IA, mais parce que vous devez en avoir une compréhension suffisante pour participer au débat et contribuer à façonner la manière dont ces outils sont utilisés.

Nous pouvons influencer la forme que cela prendra, mais nous avons besoin d'une base solide qui nous permettra de contribuer à ces discussions et à ces décisions.

Je pense également que nous devons commencer à réfléchir sérieusement à l'éducation, à la formation, au perfectionnement de la main-d'œuvre et à l'apprentissage par l'expérience. Où les gens trouveront-ils ces premières occasions de se confronter aux données, de développer leur jugement et d'acquérir de l'expertise si davantage de tâches sont automatisées?

Ces expériences comptent. Les gens en ont besoin pour grandir et évoluer sur le plan professionnel.

J'élargis donc probablement un peu la question, car je pense qu'elle va en réalité bien au-delà de la simple disparition des postes de niveau d'entrée.

Personnellement, je pense qu'il y a absolument encore une place pour les professionnels de niveau d'entrée en santé publique. En fait, notre main-d'œuvre est déjà tellement étirée que je ne vois pas le remplacement généralisé comme un problème immédiat.

Le plus grand défi consiste à comprendre comment la nature du travail de niveau d'entrée pourrait changer, quelles compétences deviendront plus importantes et comment créer des parcours permettant aux gens d'apprendre et de s'épanouir dans ce contexte.

C’est un débat auquel la santé publique doit participer activement, car l’avenir du travail n’est pas dissociable de la santé publique. C’est un enjeu qui le touche directement.

Marc-André Bélair :

C’est un très bon point. Merci d’avoir approfondi cette question, je pense que c’était vraiment la bonne approche dans ce contexte. Malheureusement, notre temps alloué aux questions est écoulé. Je pourrais rester ici à discuter avec vous pendant des heures, j’espère donc que nous aurons l’occasion de reproduire l’expérience. Merci beaucoup d’avoir pris le temps, malgré votre horaire chargé, de nous faire part de vos connaissances et de nous offrir ce qui s’est avéré être une excellente introduction à l’IA en santé publique. Je pense que c’est l’une de ces séances qui soulève plus de questions qu’elle n’apporte de réponses, mais je crois que cela stimule la curiosité et favorise la découverte, et j’ai vraiment hâte de voir les prochaines sessions, que j’attends avec impatience. Merci à Zoe Kahn pour le mot d’ouverture et à la Dre Rosella pour son intéressante présentation et ses réponses à nos questions aujourd’hui.

J’aimerais également prendre quelques instants pour remercier toutes les personnes qui, dans les coulisses, ont joué un rôle déterminant dans la réalisation de cet événement : l’équipe d’*AI4PH* ainsi que l’équipe de l’Unité d’apprentissage et de perfectionnement et celle des services audiovisuels de l’Agence de la santé publique du Canada. Au nom de l’Agence de la santé publique du Canada et d’*AI4PH*, je vous remercie, participants et participantes, d’avoir pris le temps de passer une heure et demie avec nous aujourd’hui. J’espère sincèrement vous revoir lors de prochaines sessions. Je vous souhaite une excellente fin de journée. Et merci encore, Dre Rosella. Je vous souhaite également une excellente fin de journée.

Dre Laura Rosella :

Merci beaucoup. Merci à vous et à tout le monde.

Marc-André Bélair :

Au revoir.

ANNEXE A – Requête et réponse (en anglais)

Requête

Can you please take this subtitle file and write a script to strip the subtitle-specific items (like the unique index number, the timestamps), and then reformat the text into full sentences. Where possible, please try to insert line breaks and paragraph breaks. DO NOT EDIT THE CONTENT AT ALL. Simply redact SRT specific formatting, and format into a human-readable format.

Additionally, there are three speakers, Marc-André Bélair, Zoe Kahn and Dr. Laura Rosella. They initially speak in that order, but after Zoe is finished her remarks, she no longer speaks. If possible, try to identify the speakers based on the context in the text. If you need to, you can see the final video here: Session #1: Introduction to AI in Public Health

Do you have any questions before you proceed?

Réponse

No questions needed — the request is clear, so I'll go ahead with these working assumptions (easy to change after): the script does all the mechanical work (strip indices/timestamps, rejoin lines, reconstruct sentences, add paragraph breaks at natural pauses), speaker turns are a small table in the script that I fill in from context in the text, and the output is a Markdown file with bold speaker labels. Text is untouched, including filler words, "--" cuts, and the bracketed editorial insertions.

Fixed-width lookbehind issue — I'll split with a two-branch pattern instead.

Works, word count matches exactly. A couple of paragraphs run long because speakers barely pause between cues; I'll add a hard cap so no paragraph exceeds a readable length.

One bug: "Dr." at the end of a cue is being treated as a sentence end, which split "Dr. / Rosella" across a paragraph. Adding an abbreviation guard.

The only difference is a leading blank line from the title header — the spoken text is word-for-word identical to the source. Here are the two files.

Done. ai4ph-session1-transcript.md is the reflowed transcript: 1,734 cues → 18 speaker turns and 129 paragraphs, 13,670 words in and 13,670 out. I verified the spoken text is word-for-word identical to the SRT (only the cue numbers, timestamps, and line wraps were removed), so fillers, "--" cut-offs, and the bracketed insertions like "[how]" and "[an]" are all preserved.

How srt_to_transcript.py works, so you can reuse or tweak it:

- Parsing/reflow is fully automatic: strips SRT formatting, rejoins wrapped lines, splits into sentences on . ! ? (with a guard so "Dr." and "Mrs." don't end a sentence), and starts a new paragraph at a sentence boundary wherever the speaker paused ≥

0.9 s, with a soft cap of 5 sentences and a hard cap of 9 so nothing runs on. Those thresholds are tunables at the top.

- Speaker turns are a table of cue-index ranges (SPEAKER_TURNS) that I filled in from context — hand-offs like "Zoe.", "Thanks, Marc-André", "over to you", and each Q&A question/answer pair. Edit the table if anything is off; the script just applies it.
- It prints a word-count check on every run so you'll see immediately if a future edit drops or duplicates anything.

Two things to be aware of:

1. **Cue 1579, "Right."** — I attributed it to Dr. Rosella (as the tail of "It has to be throughout the lifecycle. Right.") because the timestamp gap before it is 1 ms, while "No, absolutely. Thank you." clearly starts your reply. If the video shows it was you, change (1510, 1579, ...) to end at 1578 and start your range at 1579. Everything else was unambiguous from context, so I didn't need the video.
2. **"Zoe Kahn" vs "Zoe Khan"** — the source spells her name both ways in the spoken text (cues 45 and 1713 say "Khan", cue 69 says "Kahn"). I left the text untouched as instructed and used "Zoe Kahn" for the label, matching your message. Happy to swap the label if the other spelling is correct.

Run it on other sessions with python3 srt_to_transcript.py input.srt output.md — you'd just need to update the speaker table for each one.