



From Big Data to Big Brains: the Path to Smarter Drug Design in Canada

Mégadonnées, mégatalents : la voie à suivre pour une conception plus intelligente des médicaments au Canada

By Josh Pottel, CEO, Molecular Forecaster Inc. (MFI)

Par Josh Pottel, directeur général de Molecular Forecaster Inc. (MFI)

EVERY DAY, WE HEAR about how artificial intelligence is ready to revolutionize Canadian life sciences. Claims that AI can or will make most things 'faster, cheaper, better' are already ubiquitous. And while we are seeing AI's power in being able to analyze vast datasets—in minutes, not months—to pinpoint potential drug candidates or disease markers, there's an important question we need to ask: is AI intelligent enough to deliver on all that's being promised?

THE LIMITS OF LEARNING BY EXPOSURE

Many current AI approaches, particularly in machine learning, are designed to learn from exposure to the experience contained in vast amounts of data. Feed an algorithm enough examples of successful drugs, failed drugs, molecular structures, and bioactivity data, and it becomes adept at spotting correlations and predicting outcomes for similar new inputs.

But drug discovery is inherently messy and complex. Biological systems are intricate; the path from molecule to medicine is rarely linear; and the data that scientists are working with is often sparse, heterogeneous, and biased. So, relying solely on trends observed in historical data has significant drawbacks:

1. **"Black Box" problem:** Machine learning models can predict outcomes but not explain them, making their predictions hard to trust, especially in new chemical or biological scenarios.
2. **Sensitivity to data bias and the struggle with novelty:** AI tends to perpetuate

LA TENDANCE ACTUELLE est de considérer que l'intelligence artificielle est sur le point de révolutionner les sciences de la vie au Canada. L'on répète à l'envi que grâce à l'IA presque tout devrait être « plus rapide, moins cher et mieux ». On constate en effet que l'IA permet l'analyse de vastes ensembles de données—en quelques minutes et non en quelques mois—aux fins du repérage des candidats-médicaments et des éventuels marqueurs de maladies, mais il convient de se demander si l'IA est suffisamment puissante pour tenir toutes ses promesses.

LES LIMITES DE L'APPRENTISSAGE PAR ASSIMILATION

Nombre d'approches actuelles de l'IA, en particulier dans le domaine de l'apprentissage automatique, reposent sur l'assimilation de vastes quantités de données. Si l'on alimente un algorithme avec suffisamment d'exemples de

médicaments qui ont fait leurs preuves, de médicaments inefficaces, de structures moléculaires et de bases de données de bioactivité, il acquiert la capacité de repérer les corrélations et de prévoir les résultats de nouveaux intrants similaires.

Cependant, la recherche pharmaceutique est par nature un processus délicat et hasardeux.

Les systèmes biologiques étant complexes, il est rare que la découverte d'une molécule mène directement à un médicament. Qui plus est, les données à partir desquelles travaillent les scientifiques sont souvent éparses, hétérogènes et biaisées. Il n'est donc pas suffisant de se



Josh Pottel

To unlock the next level of AI-driven innovation in drug design, AI needs genuine “intelligence,” rooted in the fundamental meaning of the word: the ability to learn, understand, and apply knowledge.

biases in historical data, potentially overlooking innovative solutions outside familiar patterns. And, while adept at using existing data, AI often falters in creating novel molecular designs or addressing unprecedented biological challenges.

3. **Lack of foundational understanding:** AI doesn't truly understand scientific principles, recognizing functional group interactions based only on data patterns, not underlying chemistry.

In all the cases above, AI isn't giving us true intelligence in the way a human scientist can. Instead, it's sophisticated mimicry based on statistical correlation—powerful, sure, but fundamentally limited.

REDEFINING AI INTELLIGENCE

Imagine giving an 11-year-old a pile of data, expecting them to understand it and make recommendations to help you make a drug? That kid represents AI in drug design today: able to see patterns but lacking the intelligence to make sense of it. At the Conscience Symposium on Open Drug Discovery in Montreal earlier this year, renowned medicinal chemist and blogger, Derek Lowe, framed this as AI being able to give us answers, but not the reasoning behind them.

It's time to stop showing AI millions of examples of things, hoping it intuitively rules. To unlock the next level of AI-driven innovation in drug design, AI needs genuine “intelligence,” rooted in the fundamental meaning of the word: the ability to learn, understand, and apply knowledge.

TEACHING FOUNDATIONAL PRINCIPLES

Before asking an AI to design a drug, we need to teach it the basics, much like a human student, and make it smarter. This doesn't mean feeding the AI raw data points that represent the principles of chemistry, physics, and biology; it means embedding the principles themselves into the AI's architecture or knowledge base.

Think of it like giving the AI the textbooks and making sure it understands the core concepts before it sees specific experimental results. The goal is an AI that possesses a rudimentary, rule-based understanding of how the world works at a molecular and biological level.

contenter des tendances observées dans les données antérieures :

1. **Le principe de la « boîte noire » :** Les modèles d'apprentissage automatique peuvent prévoir les résultats, mais ne savent pas les justifier, ce qui fait qu'il est difficile de se fier à leurs prévisions, en particulier dans le cas de nouveaux éléments chimiques ou mécanismes biologiques.
2. **Une vulnérabilité aux biais dans les données et la difficulté à faire du neuf :** L'IA a tendance à perpétuer les biais des données passées et néglige potentiellement les solutions innovantes qui sortent des schémas habituels. Si elle est en mesure d'assimiler les données existantes, elle est souvent incapable de créer de nouveaux modèles moléculaires et de répondre à de nouveaux besoins biologiques.
3. **Un manque de compréhension des principes fondamentaux :** L'IA ignore les principes scientifiques; elle reconnaît les interactions entre les groupes fonctionnels d'après les seuls modèles de données, et non selon les fondements de la chimie.

Ainsi, la plupart du temps, l'IA n'offre pas la véritable intelligence du scientifique. Elle fonctionne par imitation sur le principe de la corrélation statistique, certes de façon assez élaborée et efficace, mais sa portée est limitée.

REDÉFINIR L'INTELLIGENCE ARTIFICIELLE

Imaginez que vous donniez à un enfant de 11 ans une série de données et que vous vous attendiez qu'il puisse formuler des recommandations pour vous aider à élaborer un médicament. L'IA agit en quelque sorte comme cet enfant relativement à la conception de médicaments : elle peut déceler des schémas, mais elle ne peut en comprendre le sens global. Lors du colloque Conscience sur la découverte de médicaments en science ouverte qui s'est tenu à Montréal au début de l'année, Derek Lowe, chimiste médical et blogueur de renom, a expliqué que l'IA pouvait proposer des solutions, mais pas le raisonnement qui les sous-tend.

Il est temps d'arrêter de transmettre à l'IA des millions de modèles en espérant qu'elle en tire des règles. Pour franchir une nouvelle étape en matière de conception de médicaments, l'IA devra faire preuve d'une véritable « intelligence », dans le sens fondamental du terme, soit une capacité à apprendre, à comprendre et à appliquer des connaissances.

L'APPRENTISSAGE DES PRINCIPES FONDAMENTAUX

S'il est envisageable de penser qu'une IA puisse concevoir des médicaments, cela passera d'abord par le fait de lui enseigner les bases, comme à un étudiant. C'est à cette condition qu'elle acquerra une véritable « intelligence ». Il ne s'agit pas d'alimenter l'IA en données brutes qui représenteraient les principes de la chimie, de la physique et de la biologie, mais d'intégrer les principes eux-mêmes dans l'architecture ou la base de connaissances de l'IA.

Intelligent AI could help us design novel chemical entities for challenging targets, predicting drug properties with higher accuracy and interpretability, and reducing late-stage failures by identifying issues earlier based on mechanistic understanding.

DEVELOPING CRITICAL THINKING AND LEARNING HOW TO SOLVE PROBLEMS

Once the AI has foundational scientific understanding, we need to teach it to apply these principles, integrating its foundational knowledge with real-world experimental data. Here, the AI learns to:

- **Reason**, moving from correlation to understanding causation based on scientific mechanisms.
- **Hypothesize**, generating molecular ideas grounded in chemical and biological plausibility.
- **Critically evaluate**, assessing drug candidates with a comprehensive understanding of their potential benefits and risks.
- **Adapt**, using fundamental rules to handle sparse or conflicting data for more robust predictions.

Once it's learned these things, the AI can move beyond replicating patterns and instead can engage in computational reasoning. For example, it can make smarter decisions in computational chemistry because it understands the underlying science. It can suggest genuinely novel molecules because its creativity isn't constrained solely by the limits of the training data.

IMPLICATIONS FOR CANADIAN BIOTECH

Canada's biotech sector is ready for truly intelligent AI in the creative and complex field of drug design.

Intelligent AI could help us design novel chemical entities for challenging targets, predicting drug properties with higher accuracy and interpretability, and reducing late-stage failures by identifying issues earlier based on mechanistic understanding. It could help us develop tools that interpret complex biomarker signatures based on an understanding of underlying disease biology, not just statistical patterns. And imagine the promise of AI that could create tailored therapeutic strategies by simulating how drug candidates might interact with an individual's specific biological makeup, informed by both genomic data and foundational biological principles.

Ce serait comme de demander à l'IA de se familiariser avec les manuels scolaires pour s'assurer qu'elle comprend les concepts de base afin qu'elle puisse ensuite passer à la phase expérimentale. Il s'agit de créer une IA qui possède une compréhension rudimentaire, basée sur des principes, de la manière dont le monde fonctionne au sens moléculaire et biologique du terme.

ACQUÉRIR UNE PENSÉE CRITIQUE ET APPRENDRE À RÉSOUDRE DES PROBLÈMES

Après cette première étape d'acquisition des connaissances scientifiques fondamentales, l'IA devrait apprendre à appliquer les principes, en intégrant ses connaissances fondamentales aux données expérimentales issues du monde réel. Pour ce faire, l'IA apprendrait à :

- **Raisonner**, c'est-à-dire passer de la corrélation à la compréhension de la causalité à partir de mécanismes scientifiques.
- **Émettre des hypothèses**, c'est-à-dire créer de nouvelles idées relativement aux molécules en se basant sur la vraisemblance chimique et biologique.
- **Évaluer avec un esprit critique** les candidats-médicaments en ayant une compréhension globale de leurs avantages et de leurs risques éventuels.
- **S'adapter**, c'est-à-dire appliquer des règles fondamentales pour aborder les données lacunaires ou contradictoires afin d'obtenir des prévisions plus fiables.

Une fois ces connaissances acquises par l'IA, on pourrait envisager de lui faire dépasser le stade de la reproduction de modèles pour effectuer un réel raisonnement computationnel. Par exemple, elle pourrait prendre des décisions plus judicieuses en matière de chimie computationnelle grâce à sa compréhension des fondements de la science. Elle serait aussi en mesure de concevoir des molécules réellement novatrices, car sa créativité ne serait pas limitée par les seules données d'entraînement.

QUELLES RÉPERCUSSIONS POUR LA BIOTECHNOLOGIE CANADIENNE?

Le secteur biotechnologique canadien est impatient de pouvoir compter sur une IA pleinement fonctionnelle dans le domaine complexe de la mise au point de médicaments.

Une telle IA pourrait contribuer à créer de toutes nouvelles entités chimiques pour les cibles complexes, à prévoir les propriétés des médicaments avec une précision et une interprétation supérieures et à limiter le risque d'échecs au stade final grâce à l'isolement précoce des problèmes, fondé sur la compréhension des mécanismes. Elle permettrait de concevoir des outils susceptibles d'interpréter la signature des biomarqueurs complexes en fonction de la compréhension du mécanisme biologique des maladies plutôt que des seuls modèles statistiques. Imaginez à quoi ressemblerait une IA apte à créer des stratégies thérapeutiques sur mesure en simulant la manière dont les candidats-médicaments pourraient interagir avec la structure biologique d'un individu,



Human intelligence collaborating with computational intelligence at Molecular Forecaster
Chez Molecular Forecaster, l'intelligence humaine travaille de concert avec l'intelligence computationnelle

There is no better time to support Canadian companies developing educated AI systems that truly understand science. It's also time to move beyond the hype that AI will make things faster, cheaper, better, and instead focus on turning AI into an educated partner that can help us develop smarter, more innovative solutions. We have a shared opportunity and responsibility to work together, now more than ever, to grow Canada's life sciences ecosystem. This is about changing the game completely to position Canada at the forefront of a new era in computationally driven biomedical discovery, ready to tackle the toughest challenges in human health. 

Josh Pottel is CEO at Molecular Forecaster Inc. (MFI), where he and his team have been teaching chemistry to their proprietary software for years, helping organizations make smarter decisions in drug design. MFI's goal? To become the go-to partner in small molecule drug design, combining proprietary tools and technology, deep expertise and know-how, and an all-in approach to collaboration that sets MFI's partners up for success.

grâce aux données génomiques et aux principes biologiques fondamentaux.

L'heure est venue d'accompagner les entreprises canadiennes qui se spécialisent dans la mise au point de tels systèmes d'intelligence artificielle. On doit également faire fi du battage médiatique qui fait de l'IA une panacée et se concentrer plutôt sur la façon de transformer l'IA en un partenaire d'apprentissage pouvant aider à élaborer des solutions vraiment utiles et novatrices. Plus que jamais, il est de notre responsabilité de travailler ensemble à la croissance de l'écosystème des sciences de la vie au Canada. L'objectif est de changer complètement la donne afin de positionner le Canada à l'avant-garde d'une nouvelle ère de découvertes biomédicales informatisées, et de le préparer à relever les défis les plus complexes dans le domaine de la santé humaine. 

Josh Pottel est le directeur général de Molecular Forecaster Inc. (MFI), une entreprise qui travaille depuis des années à « apprendre » la chimie à un logiciel exclusif, afin d'aider les organisations à prendre des décisions plus judicieuses en matière de conception de médicaments. L'objectif de MFI : devenir le partenaire de référence dans la mise au point de médicaments à petites molécules, en combinant : des outils et des technologies exclusifs; une expertise et un savoir-faire approfondis; une approche profondément collaborative qui place ses partenaires sur la voie de la réussite.