LA BIOTECHNOLOGIE DE LA BOÎTE NOIRE

Intégration de l'intelligence artificielle à la biologie de synthèse : S'attaquer aux risques, au battage médiatique et aux inégalités qui sous-tendent la biologie générative

SEPTEMBRE 2024

AFRICAN CENTRE FOR BIODIVERSITY

etc GROUP

TWN Third World Network

Le Centre Africain pour la Biodiversité (African Centre for Biodiversity, ACB) continue à s'engager à démanteler les inégalités dans les systèmes alimentaires et agricoles sur le continent africain ; il poursuit également son engagement en faveur du droit des peuples à une alimentation saine et culturellement appropriée, produite au moyen de méthodes saines et durables sur le plan écologique, et de leur droit à définir leurs propres systèmes alimentaires et agricoles.

AFRICAN CENTRE FOR BIODIVERSITY

© The African Centre for Biodiversity
www.acbio.org.za
PO Box 29170, Melville 2109, Johannesburg, South Africa.
Tel: +27 (0)11 486-1156

CC BY NC ND

Recherches et rédaction de Jim Thomas, Scan the Horizon, anciennement appelé ETC Group

Évaluation externe par Maywa Montenegro de Wit, Professeur adjoint au Service des études environnementales de l'Université de Californie, Santa Cruz, qui a nécessité des contributions substantielles, y compris celles de Dan McQuillan, chargé de cours en informatique créative et sociale au Goldsmiths College, Royaume-Uni

Directives éditoriales de la directrice générale de l'ACB Mariam Mayet

Révision en anglais de Liz Sparg
Traduction : IEBE
Conception et mise en page : Xealos

Remerciements
Nous tenons à remercier ETC Group et le Réseau Tiers-monde d'avoir examiné le texte et collaboré avec l'ACB sur les questions soulevées dans cette note d'information. L'ACB tient à remercier plusieurs donateurs pour leur soutien financier. Les opinions exprimées dans ce document ne reflètent pas nécessairement celles de nos donateurs.

Table des matières

Abréviations 4

Sommaire 5

L'enjeu 6
Incidences politiques 7
Comment nous en sommes arrivés là : Contexte de la Convention sur la diversité biologique (CDB) 7
Les connaissances de base : Que sont la « biologie de synthèse » et l'« intelligence artificielle » (IA) ? 12

Comment l'IA et la biologie de synthèse convergent-elles ? 14

1. Bioconception : Biologie générative 15
Le génie génétique devient une « boîte noire » (et échappe à la surveillance de la CDB) 18
Quelles parties des systèmes génétiques peuvent désormais être bioconçues par l'IA ? 20
2. Bioproduction : stimuler la biotechnologie 21
3. Systèmes bionumériques : l'exemple des cultures « robotisées » 22
4. Biocalcul : l'exemple des organoïdes cérébraux 22

La ruée vers la biologie générative : les titans de la technologie se lancent dans la biologie (et la politique de la biotechnologie/biodiversité) 23

Implications pour la CDB : cinq questions urgentes 26

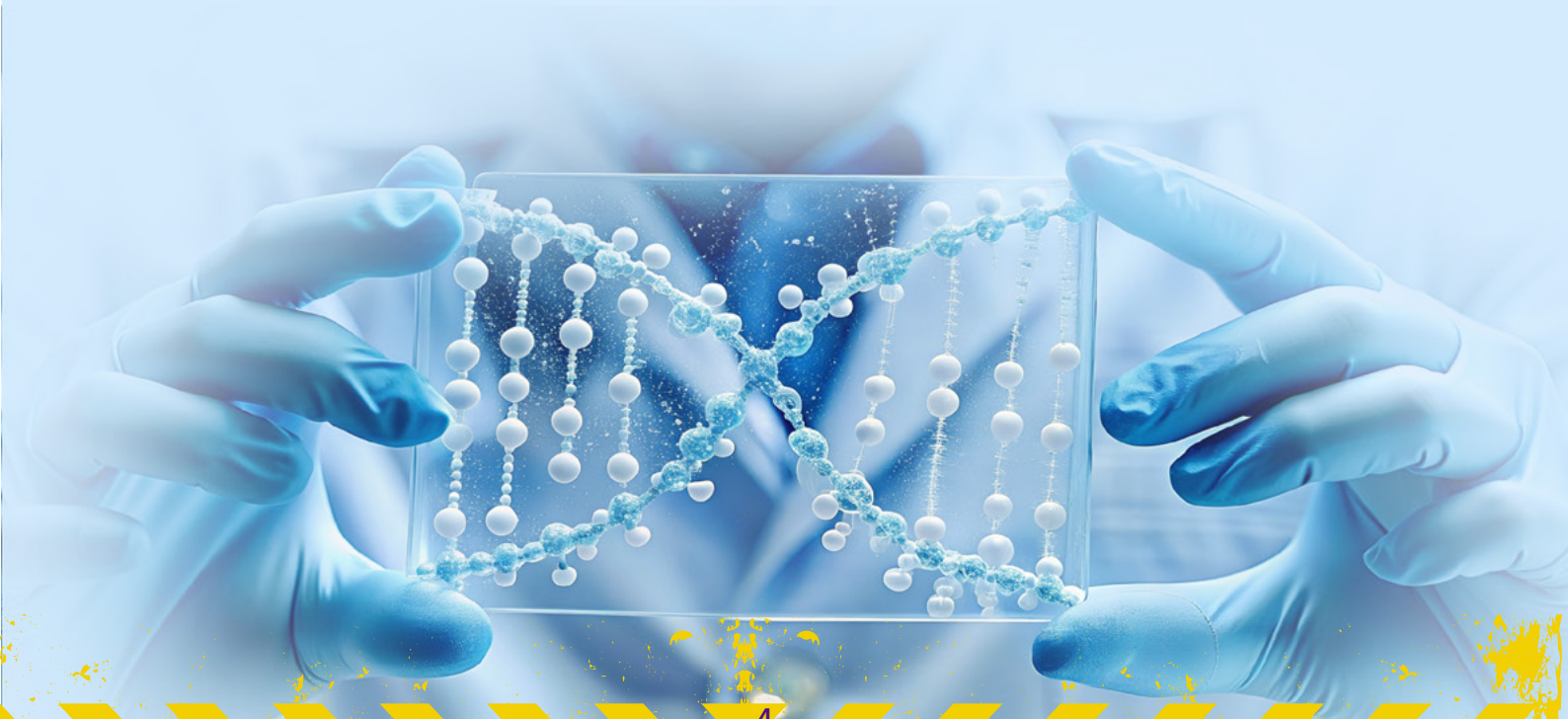
1. La biologie générative de l'IA sape-t-elle les modalités d'accès et de partage des bénéfices du Protocole de Nagoya et la gouvernance des ISN ? 26
2. La biologie générative de l'IA sape-t-elle les dispositions de biosécurité du Protocole de Cartagena sur la biosécurité ? 27
3. La biologie générative de l'IA pose-t-elle des risques en matière de biosécurité et d'armes biologiques ? 29
4. L'intégration de l'IA à la biologie de synthèse améliorera-t-elle ou détériorera-t-elle la santé et la durabilité ? 29
5. Quelles sont les implications de l'intégration de l'IA à la biologie de synthèse pour les connaissances et les pratiques traditionnelles ? 31

Une piste à suivre pour la CDB 32

Glossaire des termes clés 34

Abréviations

AHTEG	Groupe spécial d'experts techniques
CDB	Convention sur la diversité biologique
COP	Conférence des parties
GURTs	Technologie de restriction de l'utilisation des ressources génétiques
IA	Intelligence artificielle
ISN	Informations de séquençage numérique
LLM	Grand modèle de langage
mAHTEG	Groupe d'experts techniques ad hoc pluridisciplinaire (sur la biologie de synthèse)
OGM	Organismes génétiquement modifiés
SBSTTA	Organe subsidiaire chargé de fournir des avis scientifiques, techniques et technologiques
SynBio	Biologie de synthèse



Sommaire

La Convention des Nations Unies sur la diversité biologique (CDB) et ses protocoles sont le principal instrument mondial de surveillance de la biotechnologie moderne. Ils ont réussi à s'adapter au fur et à mesure qu'il y a eu des nouveautés dans le domaine. Les outils d'intelligence artificielle (IA) « génératifs » d'aujourd'hui, mieux connus pour les chatbots textuels tels que ChatGPT, sont maintenant utilisés pour générer de nouvelles séquences numériques pour les organismes génétiquement modifiés (OGM) et les protéines. Ces modèles, développés par de grandes entreprises de technologie numérique, sont entraînés sur de grandes quantités de séquences d'ADN ou de protéines numériques, trouvent des modèles et les appliquent pour créer de nouvelles séquences numériques. Cette nouvelle industrie, surnommée « biologie générative » par ses défenseurs, s'accompagne de promesses selon lesquelles de tels outils de « bioconception » de l'IA peuvent fournir un éventail de solutions technologiques pour un monde plus durable. Les revendications actuelles en faveur de la biologie générative soutiennent les spéculations faites sur les cycles précédents d'OGM et de systèmes d'IA de première génération. Chacun d'entre eux n'a pas répondu au battage publicitaire initial lorsque de nouveaux problèmes ont émergé.

Au-delà du battage médiatique, le domaine de la biologie générative représente une prise audacieuse de l'information de séquençage numérique (ISN) mondiale sur les ressources génétiques. Que des bioproduits fiables voient le jour ou non, nous assistons déjà à des investissements importants dans ces développements et à de puissants acteurs numériques qui alimentent le battage médiatique pour générer de la fascination, de l'espoir et des investissements dans la biologie générative. Avec la légitimité accordée par les bailleurs de fonds de la Silicon Valley, les entreprises d'IA tenteront probablement de changer considérablement les conditions de gouvernance de la biotechnologie moderne — affirmant que l'approche de base de la CDB, qui consiste à défendre la prudence et l'équité, est désormais dépassée à l'ère de l'IA. La CDB doit être prête à séparer la réalité de la fiction. Elle doit redoubler d'efforts de vérification sur le terrain en matière d'analyse prospective, d'évaluation des technologies et de surveillance en mettant en œuvre un processus raisonnable pour comprendre les implications de l'intégration de l'IA à la biologie de synthèse. Les parties à la CDB devraient examiner et renforcer les dispositions de la Convention en matière de surveillance de la biotechnologie face à l'évolution rapide du paysage technologique.

COP15, Biodiversité de l'ONU - **© Wikimedia Commons**

L'enjeu

L'application des outils d'IA au génie génétique et à la biotechnologie implique des changements importants qui sont examinés plus attentivement dans ce document d'information :

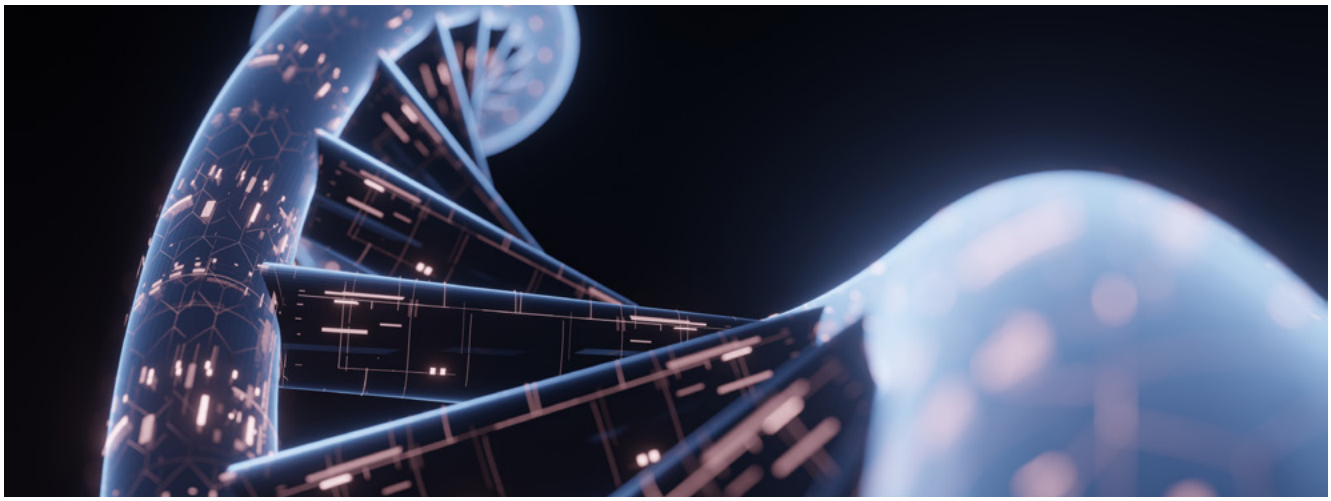
a) Biotechnologie de la « boîte noire » : L'emploi de l'IA pour concevoir numériquement des systèmes génétiques fait passer le processus de génie génétique dans une « boîte noire » algorithmique imprévisible où les décisions de conception individuelles ne peuvent être ni tracées ni expliquées. Ce caractère opaque de « boîte noire » de la bioconception de l'IA, qui est décrit plus en détail ci-dessous, est inhérent à l'IA générative. Dans le contexte de la biotechnologie, elle remet en question les capacités actuelles d'évaluation de la biosécurité, sape les exigences de surveillance et supprime la traçabilité nécessaire pour assurer un partage juste et équitable des avantages découlant de l'utilisation des ressources génétiques ou pour soutenir les systèmes de responsabilité et de réparation. Des dispositions déjà difficiles pour retracer l'origine des ISN sont sur le point de devenir malicieusement difficiles ou impossibles à la suite de l'utilisation d'outils d'IA.

b) Biotechnologie renforcée : L'industrie soutient que laisser les robots d'IA concevoir des formes de vie, des protéines et des parties génétiques (biologie générative) automatise et accélère la conception de nouveaux organismes — qui n'ont peut-être jamais existé auparavant dans la nature. Cette approche vise à accroître l'efficacité des systèmes de production biotechnologique et, si elle est efficace, elle pourrait augmenter considérablement le nombre total de nouvelles formes de vie et de protéines produites, déclenchant ainsi un déluge de nouvelles entités biotechnologiques entrant sur le marché ou dans les écosystèmes. Cela rappelle l'urgence d'une capacité humaine pour réglementer, surveiller et entreprendre une évaluation de la biosécurité. Cette situation soulève également de graves menaces pour la biosécurité.

c) Risques bionumériques : L'une des caractéristiques du développement technologique actuel est l'essor des systèmes cyberphysiques, c'est-à-dire des systèmes plus complexes du monde réel contrôlés par des technologies numériques, notamment l'IA, les capteurs et l'automatisation. Ces systèmes cyberphysiques incluent de plus en plus d'organismes, de composants et de produits de biologie de synthèse, brouillant ainsi la frontière entre les domaines biologique et numérique (ce qu'on appelle la convergence bionumérique). Ces systèmes hybrides « bionumériques », composés en partie d'organismes ou de composants génétiquement modifiés, seront de plus en plus conçus, gérés ou surveillés par l'IA, ce qui retirera les humains de la boucle et donnera lieu à de nouveaux risques et vulnérabilités. Les partisans espèrent qu'à l'avenir, les systèmes d'IA eux-mêmes pourront être basés sur l'utilisation de bio-ordinateurs génétiquement modifiés. Si cela devient un jour possible, cela entraînera des risques supplémentaires en matière de biosécurité et des enjeux de bioéthiques au cœur des industries numériques.

Incidences politiques

Les gouvernements s'efforcent déjà de rattraper les effets secondaires, les erreurs et les énigmes de gouvernance créés par les programmes d'IA générative de première génération — tels que ChatGPT — tout en découvrant la portée excessive des affirmations initialement faites par les développeurs d'IA. La CDB, forte de trois décennies d'expérience dans le suivi de la politique mondiale en matière de biotechnologie, est particulièrement bien placée pour évaluer le domaine désormais émergent de la « biologie générative » et offrir des conseils judicieux avant que les risques de l'IA ne soient irrévocablement liés aux risques de génie génétique. Lors de la seizième Conférence des Parties (COP 16) de la CDB en octobre et novembre 2024, les gouvernements auront l'occasion de commander une évaluation « plus approfondie » afin de mieux comprendre l'ensemble des défis politiques découlant de l'intégration rapide de l'IA à la biologie de synthèse — et de proposer comment relever rapidement ces défis dans le cadre de la précaution et de la justice. Les accords conclus lors de la COP 16 sur les ISN doivent également garantir que les géants de l'IA qui accumulent actuellement des ISN pour former des modèles de biologie générative sont fermement couverts par les exigences relatives à l'utilisation commerciale des ISN.



Biologie synthétique - **© kkssr, Shutterstock**

Comment nous en sommes arrivés là : Contexte de la Convention sur la diversité biologique (CDB)

Depuis trois décennies, les parties à la CDB ont pris des mesures audacieuses et préventives pour tenir la Convention et ses protocoles à jour étant donné l'évolution rapide de la biotechnologie moderne. Au départ, la CDB a été créée en tant que première culture commerciale d'OGM disséminée dans les pays développés.[1] Des décisions ultérieures de la CDB ont suivi et apporté des réponses aux nouveautés en matière de biotechnologie au fur et à mesure qu'elles se sont produites — y compris la gestion des mouvements transfrontaliers et du commerce mondial des OGM, la technologie de restriction de l'utilisation des ressources génétiques (GURTS ou « grains terminateurs »), la nécessité de prévenir la biopiraterie, l'essor de la biologie de synthèse et des organismes synthétiques, le développement du forçage génétique (une technologie qui entraîne des mutations dans les populations sauvages) et l'utilisation non réglementée des ISN (voir la figure 1). L'intégration de l'IA à la biologie de synthèse est la prochaine étape logique de l'évolution de la CDB.

1 Miller, S.K. 1994. "Genetic first upsets food lobby", *New Scientist*, 28 mai, https://www.newscientist.com/article/mg14219270-700-genetic-first-upsets-food-lobby/

FIGURE 1

Comment la CDB a réagi aux principales **avancées de la biotechnologie** au cours des 30 dernières années

G Principales avancées en matière de gouvernance dans le cadre de la CDB et des protocoles

T Principales avancées technologiques

1987 (T) La bactérie ice-minus *(Pseudomonas syringae)*, premier organisme génétiquement modifié disséminé dans l'environnement (USA)

1988 (G) Le Groupe de travail spécial d'experts sur la diversité biologique commence à négocier le texte de la CDB — englobant la réglementation de la biotechnologie moderne

1992 (G) Création de la CDB, y compris l'accord sur l'élaboration d'un protocole international de biosécurité

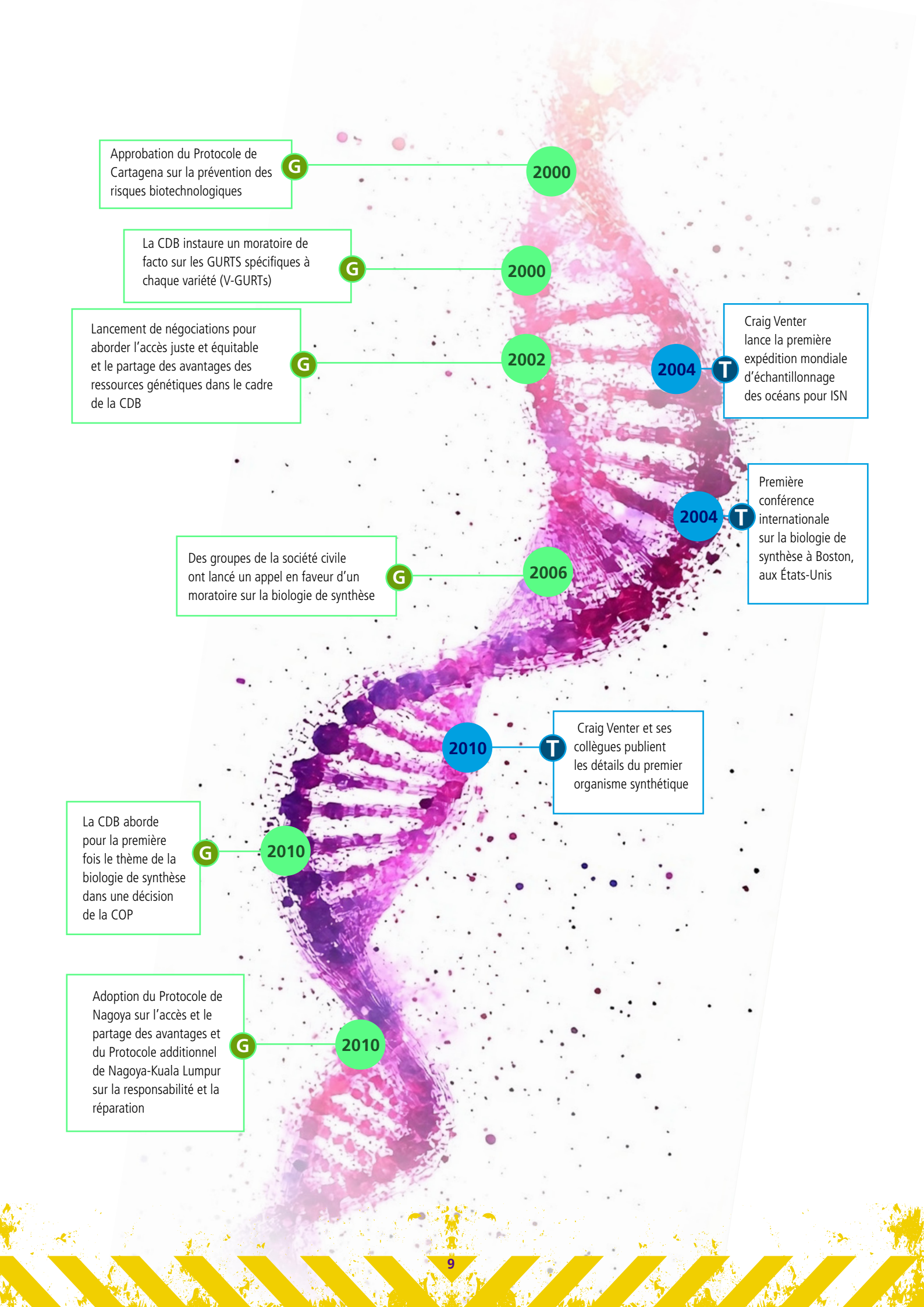
1994 (T) Première commercialisation d'une culture génétiquement modifiée par la société américaine Calgene (tomate 'Flavr Savr' CGN-89564-2)

1996 (T) Le soja tolérant aux herbicides de Monsanto devient la première culture génétiquement modifiée de base cultivée et commercialisée

1996 (G) Création du Groupe de travail de la CDB sur la prévention des risques biotechnologiques pour réglementer les mouvements transfrontaliers mondiaux d'OGM

1998 (G) La société civile découvre des recherches sur des « grains terminateur » OGM stériles (GURT)

1999 (G) La société civile publie des preuves de biopiraterie à grande échelle


Approbation du Protocole de Cartagena sur la prévention des risques biotechnologiques
G
2000
La CDB instaure un moratoire de facto sur les GURTS spécifiques à chaque variété (V-GURTs)
G
2000
Lancement de négociations pour aborder l'accès juste et équitable et le partage des avantages des ressources génétiques dans le cadre de la CDB
G
2002
2004
T
Craig Venter lance la première expédition mondiale d'échantillonnage des océans pour ISN
2004
T
Première conférence internationale sur la biologie de synthèse à Boston, aux États-Unis
Des groupes de la société civile ont lancé un appel en faveur d'un moratoire sur la biologie de synthèse
G
2006
2010
T
Craig Venter et ses collègues publient les détails du premier organisme synthétique
La CDB aborde pour la première fois le thème de la biologie de synthèse dans une décision de la COP
G
2010
Adoption du Protocole de Nagoya sur l'accès et le partage des avantages et du Protocole additionnel de Nagoya-Kuala Lumpur sur la responsabilité et la réparation
G
2010

Création du premier Groupe spécial d'experts techniques (AHTEG) de la CDB sur la biologie de synthèse

G 2014

2014 T

Premier système de forçage génétique fonctionnel (chez les mouches des fruits) signalé

L'AHTEG de la CDB sur la biologie de synthèse soulève d'abord des préoccupations concernant l'utilisation inéquitable des ISN et les risques écologiques des systèmes de forçage génétique

G 2015

La CDB établit une définition opérationnelle de la biologie de synthèse

G 2016

La CDB décide d'entreprendre des travaux sur les ISN

G 2016

Première décision de la COP de la CDB concernant le forçage génétique

G 2016

L'AHTEG de la CDB propose l'analyse prospective, l'évaluation et la surveillance des nouveaux développements dans le domaine de la biologie de synthèse — et soulève d'abord l'utilisation de l'IA en tant que tendance nouvelle et émergente

G 2017

2018 T

L'AlphaFold 1 de DeepMind remporte le concours CASP (Critical Assessment of Structural Prediction) — première utilisation majeure de l'IA dans la biologie de synthèse

La COP de la CDB s'accorde sur la nécessité d'un processus d'analyse prospective, d'évaluation et de surveillance pour la biologie de synthèse

G 2019

2022 T

Dans Science, les scientifiques mettent en garde contre le développement rapide et les risques des vaccins autopropagateurs modifiés

La CDB établit un processus d'analyse prospective, d'évaluation et de suivi des nouveaux développements dans la biologie de synthèse, y compris le Groupe d'experts techniques ad hoc multidisciplinaire (mAHTEG)

G 2022

2022 T

La sortie publique de ChatGPT2 catalyse une nouvelle ruée vers les investissements dans l'IA générative

Le mAHTEG sur la biologie de synthèse identifie l'intégration de l'IA à la biologie de synthèse (ainsi que le développement de vaccins autopropagateurs) comme des tendances clés urgentes auxquelles la CDB doit s'attaquer grâce à une évaluation plus approfondie

G 2023

IA édition de gènes - **© Rawpixel.com, Shutterstock**

À la lumière du développement rapide de nouvelles biotechnologies dans le domaine de la biologie de synthèse, les parties à la COP 14 de la CDB en 2018 ont convenu de la nécessité d'un processus large et régulier d'analyse prospective, d'évaluation et de surveillance qui pourrait identifier les principales tendances et évolutions de la biologie de synthèse suffisamment tôt pour prendre des mesures de gouvernance.[2] Un tel processus a ensuite été officiellement lancé lors de la COP 15 en 2022 pour aider la CDB à rester pertinente face aux développements en cours dans le domaine de la biologie de synthèse.[3] Un forum en ligne, suivi d'une série de réunions virtuelles et en personne du mAHTEG sur la biologie de synthèse, a permis d'identifier cinq tendances clés pour lesquelles une évaluation initiale a été entreprise. Il s'agissait notamment du thème « Intégration de l'intelligence artificielle à la biologie de synthèse ». Dans son rapport au SBSTTA 26,[4] les experts du mAHTEG ont exhorté les parties à la CDB à agir sur le thème de l'intégration de l'IA à la biologie de synthèse. Le mAHTEG a indiqué que « le développement accéléré de l'intelligence artificielle et de l'apprentissage automatique dans le domaine de la biologie de synthèse peut avoir des impacts négatifs importants sur les objectifs, les principes et les dispositions de la Convention et que ces impacts potentiels doivent être évalués plus avant ».[5]

Le mAHTEG a proposé :[6]

- La CDB entame un processus de formulation de politiques.
- « Une demande adressée au Groupe spécial pluridisciplinaire d'experts techniques pour qu'il entreprenne une évaluation plus approfondie en vue d'établir un rapport abordant, entre autres, les incidences potentielles sur la prévention des risques biotechnologiques, l'utilisation durable de la diversité biologique, l'accès équitable et le partage des avantages, les aspects sociaux, économiques et culturels, les incidences sur les connaissances et les pratiques traditionnelles, et d'autres questions pertinentes ».
- Le secrétariat de la CDB élabore une série de publications techniques sur l'IA et participe aux activités du système des Nations Unies sur l'IA.
- Les parties « envisagent l'élaboration de dispositions de gouvernance efficaces et équitables pour les ensembles de données d'intelligence artificielle, les modèles de base, les outils de bioconception algorithmique, les outils de science automatisée et l'utilisation d'organismes, de composants et de produits de biologie de synthèse dans les systèmes cyberphysiques ».

Malheureusement, lors du SBSTTA 26 de la CDB, des discussions constructives sur la biologie de synthèse ont été gravement entravées par un petit groupe de pays alignés sur l'industrie de la biotechnologie qui ont déclaré qu'ils souhaitaient mettre fin au processus d'analyse prospective, d'évaluation et de surveillance qui avait été convenu précédemment.[7] Malgré les obstructions de ce petit groupe, des termes de référence fortement crochetés ont néanmoins été rédigés, qui comprenaient une demande pour que le mAHTEG :

2 Décision CBD/COP/DEC/14/19 — 30 novembre 2018 : « *Convient* qu'une analyse prospective, un suivi et une évaluation larges et réguliers des développements technologiques les plus récents sont nécessaires pour examiner les nouvelles informations concernant les impacts positifs et négatifs potentiels de la biologie synthétique par rapport aux trois objectifs de la Convention et à ceux du Protocole de Cartagena et du Protocole de Nagoya ».

3 Décision CBD/COP/DEC/15/31 paragraphe 4

4 La 26e réunion de l'Organe subsidiaire chargé de fournir des avis scientifiques, techniques et technologiques (SBSTTA) s'est tenue à Nairobi, au Kenya, du 13 au 18 mai 2024. Cet organe fonctionne dans le cadre de la CDB et fournit à la COP des conseils sur les questions scientifiques, techniques et technologiques liées à la mise en œuvre de la Convention.

5 Rapport du Groupe d'experts techniques ad hoc pluridisciplinaire sur la biologie de synthèse dans CBD/SBSTTA/26/4 — Annexe 1

6 Ibid

7 Sirinathsinghji, Eva. 2024. « SynBio discussions held to ransom », SBSTTA 26 Daily Eco report of the CBDA, 18 mai, https://cbd-alliance.org/sites/default/files/2024-05/ECO-6-SBSTTA-26.pdf

Poursuive l'évaluation approfondie, y compris sur [leurs] impacts [positifs et négatifs] potentiels [à la lumière de] [sur] les objectifs de la Convention [écologiques, socio-économiques — y compris les menaces pour les moyens de subsistance, l'utilisation durable de la biodiversité — et les considérations éthiques et culturelles, et en tenant particulièrement compte des peuples autochtones et des communautés locales, des femmes et des jeunes] de : (i) Les impacts potentiels de l'intégration de l'intelligence artificielle et de l'apprentissage automatique en biologie de synthèse.[8]

Lors de la COP 16 de la CDB, les parties ont la possibilité de s'assurer qu'une telle évaluation de l'intégration de l'IA à la biologie de synthèse progresse, conformément à l'approche de précaution, et qu'elle n'est pas davantage bloquée par des intérêts industriels.

Les connaissances de base : Que sont la « biologie de synthèse » et l'« intelligence artificielle » (IA) ?

La biologie de synthèse englobe plusieurs nouveautés et développements émergents dans la biotechnologie moderne (également connue sous le nom de génie génétique). Les premiers organismes génétiquement modifiés ont été fabriqués en laboratoire à l'aide d'une bactérie *(Agrobacterium tumefaciens)* et d'enzymes de restriction pour transférer l'ADN d'un organisme dans le génome d'un autre *(une technique connue sous le nom d'ADN recombinant, ou ADNr)*. Cependant, les biotechnologistes s'appuient de plus en plus sur une famille de technologies d'édition de gènes, telles que CRISPR/Cas, qui coupent et modifient l'ADN de l'hôte à l'aide d'enzymes guidées par l'ARN combinées à différentes voies de réparation cellulaire. L'édition génomique peut être utilisée pour supprimer une seule base afin de désactiver la fonction d'un gène (une « suppression »); elle peut également consister à insérer de petites séquences ou de nouveaux gènes entiers, y compris de donneurs « transgéniques », pour corriger ou obtenir une fonction. Comme les techniques de génie génétique précédentes, ces techniques d'édition de gènes sont actuellement annoncées pour des gains sans précédent en termes de vitesse, de bon marché et de précision. Pourtant, n'importe laquelle de ces caractéristiques techniques peut exacerber les dommages environnementaux et les inégalités sociales existants. De plus, l'histoire de la biotechnologie montre que chaque génération de percées est régulièrement proclamée supérieure à la technologie « sujette aux erreurs » qu'elle supprime, mais qu'elle affiche ensuite en peu de temps ses propres erreurs. L'édition génomique, par exemple, continue de se heurter à des défis techniques, notamment des effets non intentionnels et hors cible.[9]

Alors que l'édition génomique et l'ADNr pour l'agriculture ont évolué en grande partie à l'intersection de la biologie moléculaire et de la sélection végétale et animale, la biologie de synthèse a des racines théoriques et méthodologiques dans l'ingénierie. La biologie de synthèse applique des principes d'ingénierie pour développer de nouvelles pièces, dispositifs et systèmes biologiques ou pour reconcevoir des systèmes existants trouvés dans la nature. Les biologistes synthétiques — qui couvrent les domaines de l'informatique, de la biologie des systèmes, de l'ingénierie et de la biophysique, entre autres — conçoivent de plus en plus de nouveaux codes d'ADN dans un ordinateur et fabriquent des brins d'ADN synthétiques à partir de produits chimiques. Ils cherchent également de nouveaux moyens de

8 Décision 26/4 de la SBSTTA de la CDB — CBD/SBSTTA/REC/26/4, disponible à l'adresse https://www.cbd.int/doc/recommendations/sbstta-26/sbstta-26-rec-04-en.pdf

9 "Unintended changes induced by CRISPR/CAS cause novel risks", *Testbiotech*, dernière modification le 4 décembre 2019, https://www.testbiotech.org/en/news/unintended-changes-induced-crisprcas-cause-novel-risks/

concevoir, construire et utiliser des molécules génétiques autres que l'ADN,[10] tels que l'ARN, et des protéines telles que les enzymes, ainsi que des éléments régulateurs génétiques. Selon leur formation, les biologistes synthétiques ont tendance à traiter les molécules génétiques vivantes comme s'il s'agissait d'un code logiciel programmable pour des organismes, s'efforçant de réécrire ce « code » à des fins industrielles. Cette approche réductrice prend tout son sens dans les contraintes et les conditions d'un modèle en laboratoire. Cependant, la vie ne fonctionne pas, en fait, comme un ordinateur et la métaphore de la machine échoue souvent dans les systèmes complexes du monde réel.[11]

L'intelligence artificielle (IA) est un ensemble de technologies basées principalement sur l'apprentissage automatique et l'apprentissage profond, et utilisées pour l'analyse de données, les prédictions et les prévisions, la catégorisation d'objets, le traitement du langage naturel, les recommandations, la récupération de données, etc.

L'IA traditionnelle, également connue sous le nom d'IA basée sur des règles ou d'IA déterministe, s'appuie sur des règles et des algorithmes préprogrammés pour effectuer des tâches spécifiques. Ce type d'IA existe depuis des décennies, évoluant avec les progrès de la puissance de calcul. Contrairement aux ensembles de données ouverts utilisés dans l'IA générative, les systèmes d'IA traditionnels utilisent généralement des ensembles de données structurés et organisés et des modèles adaptés à des objectifs particuliers. En tant que tels, ces systèmes d'IA sont souvent utilisés dans la finance, la fabrication et les soins de santé pour résoudre des problèmes limités et effectuer des tâches répétitives.

Les logiciels d'IA discriminante peuvent apprendre leurs propres règles. Le logiciel d'IA prend de grands ensembles de données non structurées et les trie automatiquement en regroupements significatifs ou identifie des images ou des modèles. Pour ce faire, le logiciel d'IA est d'abord « entraîné » sur d'autres données et est aidé à faire des associations jusqu'à ce qu'il développe ses propres modèles (avec des pondérations structurées) pour faire des distinctions. Ce processus d'entraînement, inspiré des réseaux neuronaux du cerveau, est appelé « apprentissage automatique ».

L'IA générative diffère sur des points clés. À l'instar de l'IA traditionnelle et de l'IA discriminante, l'IA générative s'entraîne sur de grands ensembles de données. Mais alors que l'IA traditionnelle se limite à la prédiction selon des règles et que l'IA discriminante trie les données existantes, les modèles d'IA générative utilisent leurs ensembles de données d'entraînement pour générer de nouveaux exemples de données. L'IA générative est donc capable de créer de nouveaux résultats à partir d'un espace vectoriel qu'elle a généré à partir de données apprises. Par exemple, alors qu'un système d'IA discriminante est formé pour distinguer différentes images d'animaux, on pourrait demander à une IA générative d'examiner les relations au sein de ses données pour générer une nouvelle image d'un « chat » en fonction des matériaux qu'elle a vus et du modèle qu'elle a développé. L'IA générative utilise souvent de très grandes quantités de données d'entraînement récoltées sur Internet pour synthétiser des images, générer du texte ou transférer un style.

Depuis fin 2019, date à laquelle ChatGPT-2 a été rendu public, la grande majorité de l'IA générative utilisée est basée sur des modèles de transformateurs (grands modèles de langage) et des modèles

10 Schmidt, M. 2010. "Xenobiology: A new form of life as the ultimate biosafety tool", *Bioessays*. 32(4):322-31. doi: 10.1002/bies.200900147. PMID: 20217844; PMCID: PMC2909387. https://pubmed.ncbi.nlm.nih.gov/20217844/

11 Boudry, M., et Pigliucci, M. 2013. "The mismeasure of machine: Synthetic biology and the trouble with engineering metaphors", *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, Volume 44, Issue 4, Part B

de diffusion (voir glossaire). Par exemple, le populaire programme d'IA générative ChatGPT est un grand modèle de langage basé sur un transformateur qui s'appuie sur leur ensemble de données d'entraînement massif de langage extrait d'Internet, pour assembler du texte d'apparence crédible. Des programmes comme Stable Diffusion, DALL-E et Midjourney font de même, mais pour les images (au grand dam des artistes et des designers dont les œuvres créatives sont appropriées). Ces programmes utilisent leurs règles internes pour prédire ce qui pourrait être l'image ou la chaîne de mots correcte pour satisfaire la requête en langage naturel de l'utilisateur (« prompt »). Cette prédiction peut sembler crédible, mais après une analyse plus approfondie, elle peut être fausse (et c'est souvent le cas — voir l'encadré sur les hallucinations, les biais et l'empoisonnement des données de l'IA). Les informaticiens Arvind Narayanan et Sayash Kapoor expliquent que la faillibilité des grands modèles de langage comme ChatGPT est due en partie au fait que l'IA est incapable de déterminer la causalité ou le contexte, qui donnent tous deux un sens au langage.[12] Les modèles établissent simplement des liens statistiquement probables, fonctionnant comme ce que les experts ont décrit comme un « perroquet stochastique ».[13] Contrairement à l'IA dans la science-fiction, ces systèmes n'« apprennent » pas réellement de manière comparable à l'apprentissage du cerveau humain/animal ; L'IA n'est pas délibérative, consciente, ou malgré son nom, « intelligente ». Il s'agit cependant d'un puissant moteur de prédiction basé sur une très grande quantité de données.

Jusqu'à présent, l'IA générative s'avère être plus un battage médiatique qu'une réalité. Ses cas de test comme ChatGPT démontrent que le simple fait d'ajouter de plus grandes quantités de données n'équivaut pas à une plus grande précision ou à un pouvoir prédictif. Comme tous les modèles, ces systèmes sont entravés par les limites et les données biaisées sur lesquelles ils sont entraînés (encore une fois, voir la section ci-dessous sur les hallucinations, les biais et l'empoisonnement des données de l'IA), mais aussi par des hypothèses erronées sous-jacentes à la technologie.

Comment l'IA et la biologie de **synthèse convergent-elles ?**

Le point de vue des « grandes entreprises technologiques » :

> **Les limites sont maintenant dépassées. Nous approchons un point d'inflexion avec l'arrivée de ces technologies d'ordre supérieur, les plus profondes de l'histoire. La vague technologique à venir repose principalement sur deux technologies polyvalentes capables de fonctionner aux niveaux les plus vastes et les plus granulaires : l'intelligence artificielle et la biologie de synthèse.**
>
> **– Mustafa Suleyman, directeur général de Microsoft AI, fondateur d'Infection AI et cofondateur de DeepMind (aujourd'hui propriété de Google)[14]**

12 https://press.princeton.edu/books/hardcover/9780691249131/ai-snake-oil

13 https://dl.acm.org/doi/10.1145/3442188.3445922

14 Suleyman, M. *The coming wave: Technology, power and the 21st century's greatest dilemma*. 2023. Crown Publishing Group, Random House, p.55.

Il y a au moins quatre façons dont l'IA et la biologie de synthèse s'intègrent et celles-ci sont décrites plus en détail ci-dessous :

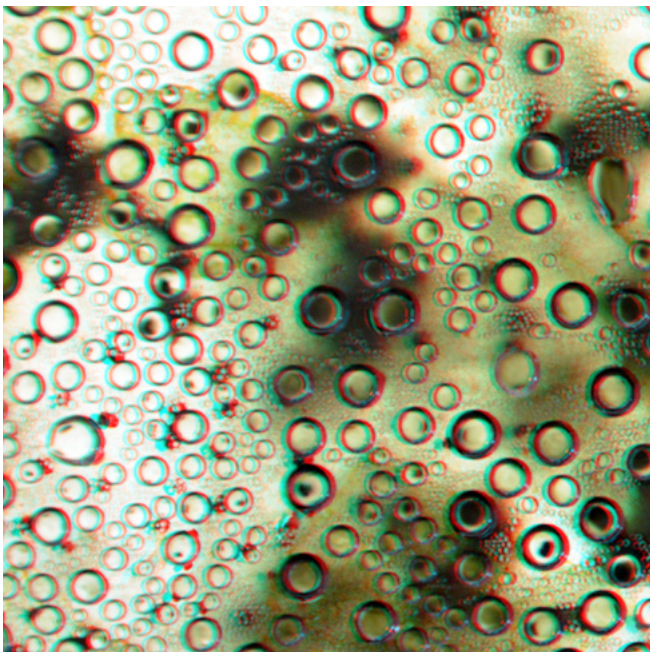
Bioconception (biologie générative) : modèles d'IA qui génèrent de nouvelles séquences génétiques ou protéiques. La conception est numérique, mais l'ingénierie des protéines et la biochimie sont utilisées pour les transformer en molécules biologiques synthétiques.

Bioproduction : L'utilisation de l'IA pour améliorer le matériel et l'efficacité des processus utilisés dans la production biotechnologique industrielle (tels que la fermentation en cuve ou la culture cellulaire dans des boîtes de Pétri).

Systèmes bionumériques (cyberphysiques) : Systèmes techniques qui mélangent des composants d'IA et de biologie de synthèse (par exemple, en utilisant des cultures modifiées ou des pulvérisations d'ARN dans l'agriculture numérique où la prise de décision est effectuée par des algorithmes).

Biocalcul : Utilisation d'ADN, de cellules, de bactéries ou d'autres parties biologiques modifiées comme matériel informatique pour les processus d'IA.

1. Bioconception : Biologie générative



Biodésign - © wim hoppenbrouwers, Flickr

Tout comme les modèles d'IA génératifs pour le texte et les images peuvent sembler créer de nouveaux écrits ou de nouvelles images en sélectionnant et en réorganisant les ensembles massifs de données sur lesquels ils ont été formés, les modèles d'IA entraînés sur de grandes quantités d'informations biologiques peuvent sembler générer de nouveaux dispositifs apparemment crédibles pour les molécules biologiques — y compris de nouvelles séquences d'ADN et de protéines. L'idée est que, plutôt que d'entraîner un LLM IA sur les textes de langues telles que l'anglais ou le français, le même modèle peut être formé sur ce que les biologistes synthétiques décrivent de manière simpliste comme des « langages » biologiques — tels que le code de l'ADN ou la séquence d'acides aminés qui composent les protéines. Le logiciel d'IA développe ensuite son propre modèle pour trier et « comprendre » ces codes biologiques, qui, à leur tour, peuvent être incités à générer un « nouveau » langage biologique — tel que de nouveaux codes d'ADN pour les génomes ou de nouveaux arrangements d'acides aminés pour les nouvelles protéines. Tous ces processus dépendent de l'acceptation d'une compréhension très réductrice (et incomplète) du fonctionnement des systèmes génétiques.

AlphaFold de Google : le « moment ChatGPT » de la biologie générative

La « preuve de concept » révolutionnaire permettant aux modèles d'IA d'interpréter la biologie et d'effectuer la biologie de synthèse était un modèle d'IA appelé AlphaFold, développé par DeepMind, la filiale d'IA de Google. AlphaFold a d'abord été formé sur des données décrivant 170 000 protéines provenant d'un référentiel public de séquences et de structures protéiques.[15]

Son objectif était de discerner comment les protéines disposées dans certaines séquences d'acides aminés se replient ensuite en structures fonctionnelles spécifiques. Au début de ce siècle, le travail scientifique nécessitait d'expliquer le lien entre le code d'une protéine unique et la manière dont elle s'est ensuite repliée dans une forme qui pourrait occuper la durée d'un doctorat entier. En 2018, AlphaFold a prédit avec succès les structures de repliement de 13 protéines à partir de leur code protéique, et en 2021, Google affirmait qu'AlphaFold 2 avait prédit les arrangements de repliement pour tous les codes de protéines connus (près de 200 millions de protéines)[16] — bien que d'autres aient observé qu'il s'agissait d'une affirmation exagérée et non étayée par la science de laboratoire.[17]

Impressionnées par ces résultats (et les communiqués de presse de Google), les publications scientifiques ont proclamé qu'AlphaFold IA avait « résolu » le problème du repliement des protéines. En mars 2024, Google DeepMind a publié AlphaFold 3, qui, selon lui, peut non seulement prédire n'importe quelle structure protéique, mais aussi « la structure et les interactions de toutes les molécules de la vie avec une précision sans précédent ».[18] Dans ses communiqués de presse, Google revendique qu'AlphaFold 3 est capable de prédire la structure de n'importe quel ADN, ARN, ligands, protéines, et plus encore, et aussi de prédire comment ces biomolécules pourraient interagir les unes avec les autres au niveau cellulaire. Après avoir formé un modèle censé être capable de faire des prédictions aussi puissantes sur les biomolécules de base de la vie, les entreprises d'IA générative traitent maintenant la prédiction comme un fait et prétendent qu'elles peuvent générer de manière fiable de nouvelles protéines, de l'ADN, de l'ARN et d'autres structures jamais vues auparavant pour le génie génétique, en utilisant la bioconception générative dirigée par l'IA. Ainsi, la « biologie générative » s'est imposée comme une nouvelle industrie. En particulier, il y a une ruée vers les nouvelles entreprises de « conception de protéines » d'IA inspirées par AlphaFold (voir ci-dessous).

15 "Method of the Year 2021: Protein structure prediction", *Nat Methods* 19, 1 (2022). https://doi.org/10.1038/s41592-021-01380-4

16 Callaway, Ewen. 2022. "'The entire protein universe': AI predicts shape of nearly every known protein". *Nature*. 608 (7921): 15–16

17 Terwilliger TC, Liebschner D, Croll TI, Williams CJ, McCoy AJ, Poon BK, Afonine PV, Oeffner RD, Richardson JS, Read RJ, Adams PD. 2024. "AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination", *Nat Methods*.

18 Google DeepMind Lab, AlphaFold Team, "AlphaFold 3 Predicts the Structure and Interactions of All of Life's Molecules", 8 mai 2024, https://blog.google/technology/ai/google-deepmind-isomorphic-alphafold-3-ai-model/#life-molecules.

Avertissement

Certes, AlphaFold en a impressionné plus d'un en tant qu'exemple de grande science pilotée par l'IA, les spécialistes des protéines travaillant en laboratoire ont déjà tiré la sonnette d'alarme pour ne pas se laisser aveugler par les capacités vantées de la biologie générative.[19] Comme tous les systèmes d'IA, AlphaFold ne résout pas le repliement des protéines en soi, il offre simplement des prédictions probables. Les scientifiques des protéines doivent encore effectuer le travail de laboratoire pour vérifier l'exactitude de ces prédictions numériques et, dans plusieurs cas et classes de protéines, les prédictions se sont avérées fausses ou trompeuses. Cela fait écho à la manière dont le battage commercial initial de l'IA générative dans le texte (ChatGPT) ou les images (Dall-E ou Midjourney) a effacé les vrais problèmes et les limites de la technologie de l'IA générative. Cela fait également écho à la manière dont les premières affirmations sur la précision de l'ADNr et de l'édition de gènes ultérieures ont été largement exagérées. Il est à craindre que, dans la ruée commerciale en cours pour utiliser des modèles d'IA générative pour générer de nouveaux systèmes génétiques vivants, les décideurs politiques ignorent les erreurs et les dangers. Contrairement au texte numérique ou aux graphiques numériques, les nouvelles formes de vie et les protéines conçues numériquement peuvent avoir un impact biologique très direct et imprévisible sur le monde vivant, surtout si elles sont mal conçues, sur la base d'hypothèses erronées et libérées de manière irresponsable.

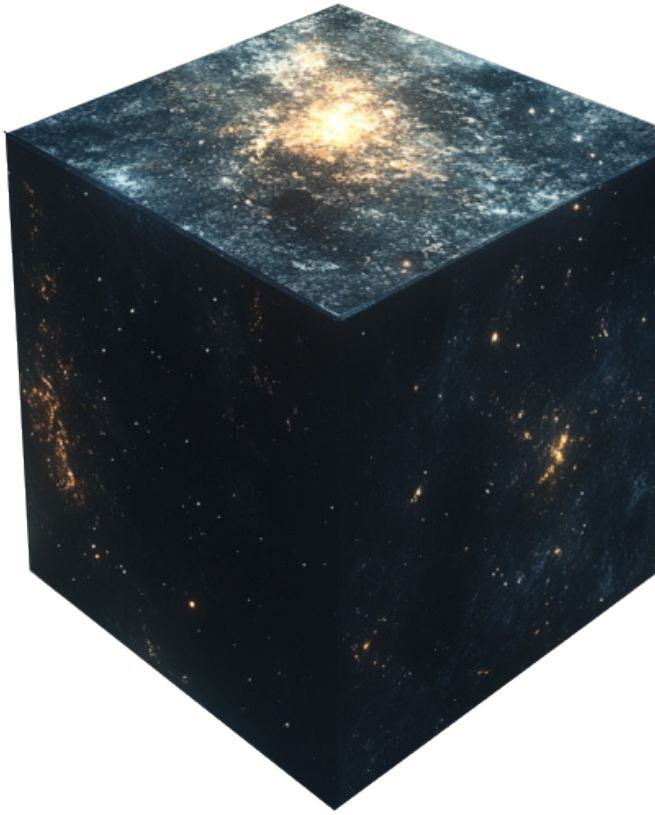
19 Voir le fil Reddit, "People are overestimating Alphafold and it's a problem", https://www.reddit.com/r/labrats/comments/1b1l68p/people_are_overestimating_alphafold_and_its_a/

Le génie génétique devient une « boîte noire » (et échappe à la surveillance de la CDB)

Une métaphore importante qui s'est développée à partir de la recherche critique sur l'IA est le concept et les risques de la « boîte noire ».[20] Le « problème de la boîte noire » fait référence à la manière opaque dont un système d'IA développe son propre modèle pendant l'« entraînement » de manière à appliquer des poids statistiques à certaines connexions qu'il observe. Le grand nombre d'opérations entreprises rend le raisonnement de ce modèle auto-généré peu transparent et inconnaissable pour un observateur humain, même un expert technique. Par conséquent, lorsqu'un système d'IA générative génère une image particulière ou fait une prédiction particulière en sortie, il n'est actuellement pas possible d'interpréter ou d'expliquer pourquoi il a généré ce résultat — par exemple, pourquoi il a choisi un mot, une image ou une séquence génétique particulière dans son résultat.

En ce qui concerne le génie génétique, l'occultation des décisions de conception biologique dans la « boîte noire » marque un changement fondamental et conséquent dans le mode d'exécution des interventions de génie génétique, en excluant l'homme de la boucle des décisions de conception. Auparavant, un ingénieur génétique humain sélectionnait ou modifiait une séquence génétique ou protéique existante, numérique ou réelle, pour une raison particulière explicable ou une conjecture qu'il pouvait étayer. Il les testait ensuite expérimentalement pour parvenir à un résultat, en documentant généralement le processus. Aujourd'hui, un système d'IA automatisé applique un modèle interne auto-généré qui ne peut être expliqué, basé sur des prédictions statistiques issues d'une analyse répétitive de milliers ou de millions de séquences. L'ordinateur sélectionne ensuite des combinaisons de séquences génétiques tirées de ses données d'apprentissage, sans en identifier la source, et les enchaîne pour former une nouvelle séquence. C'est ainsi que les décisions en matière de bioconception génétique sont transférées dans une « boîte noire » inconnaissable.

20 Voir par exemple "AI's Mysterious Black Box Problem Explained", University of Michigan-Dearborn, consulté le 25 juillet 2024, https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained.



Bien sûr, la création d'un processus de boîte noire est un choix que les concepteurs d'un système d'IA espèrent gagnant. En ce sens, ils devraient devenir directement coupables de tous les résultats de leur invention.

Les systèmes d'IA qui effectuent du génie génétique en boîte noire ont au moins trois implications importantes qui affectent la gouvernance :

1. **La biosécurité —** ISi le système d'IA qui conçoit une intervention de génie génétique ne peut expliquer ni pourquoi il a sélectionné une séquence génétique ni même à partir de quel matériel de départ cette séquence a été tirée, il devient difficile d'entreprendre une analyse indépendante de la biosécurité de cette conception. Les incertitudes associées au génie génétique augmentent considérablement, tout comme l'incapacité à repérer les problèmes, les erreurs et les fautes.

2. **L'accès et le partage des avantages —** Si le concepteur d'une intervention de génie génétique ne divulgue pas les séquences numériques dont le dessin a été tiré ou dont il a tiré (et encore moins d'où provient la source naturelle), il n'est pas possible de partager les avantages avec le détenteur initial de cette ressource génétique ou pour ce dernier d'exercer ses droits d'accès. Il devient donc impossible d'assurer un accès approprié et un partage juste et équitable des avantages.

3. **Les responsabilité et les réparations —** Si le concepteur d'une intervention génétique est un système informatique probabiliste automatisé qui génère son propre modèle opaque pour prendre des décisions, lorsque le résultat pose problème, le responsable légal n'est pas clairement établi. Il n'y a pas de contrôle humain ou de responsabilité sur le processus de génie génétique. Même si les développeurs de systèmes d'IA savent qu'ils construisent des systèmes de boîte noire qui peuvent être truffés d'erreurs, certains tribunaux ont estimé que les développeurs d'IA fournissant des logiciels pour les voitures autonomes ne sont pas responsables des accidents de voiture mortels, où un humain est techniquement un conducteur de secours.[21] Cependant, la loi est moins claire avec les voitures entièrement autonomes.[22]

Ces trois domaines (« biosécurité », « accès et partage des avantages » et « responsabilité et réparation ») décrivent les objectifs des trois protocoles de la CDB (Protocole de Carthagène, Protocole de Nagoya et Protocole de Nagoya-Kuala Lumpur) qui soutiennent la gouvernance de la biotechnologie et son impact sur le monde vivant. De plus, la CDB tente d'aborder les facteurs socio-économiques qui se poseront également lorsque les systèmes d'IA passeront à la biotechnologie. En déplaçant le génie génétique dans la boîte noire, l'intégration de l'IA à la biologie de synthèse remet directement en question l'approche de gouvernance de base que la CDB a établie dans le domaine de la biotechnologie moderne.

21 David Shepardson et Heather Somerville, "Uber Not Criminally Liable in Fatal 2018 Arizona Self-Driving Crash: Prosecutor", Reuters, 5 mars 2019, https://www.reuters.com/article/business/uber-not-criminally-liable-in-fatal-2018-arizona-self-driving-crash-prosecutor-idUSKCN1QM2P3/.

22 Jenna Greene, Column, "Driverless Car Problems Are Outpacing Liability Laws", Reuters, December 11, 2023, https://www.reuters.com/legal/transactional/column-driverless-car-problems-are-outpacing-liability-laws-2023-12-11/.

Quelles parties des systèmes génétiques peuvent désormais être bioconçues par l'IA ?

L'ADN et l'ARN — Les LLM d'IA entraînés sur toutes les ISN existantes dans le monde peuvent générer de nouvelles séquences d'ADN ou d'ARN pour les applications de biologie de synthèse.

> **Nous avons pris toutes les données ADN disponibles — à la fois les données ADN et les données d'ARN pour les virus et les bactéries que nous connaissons — environ cent dix millions de génomes de ce type. Nous avons appris un modèle de langage à ce sujet et nous pouvons maintenant lui demander de générer de nouveaux génomes.**
>
> **– Anima Anandkumar de NVIDIA Corporation décrivant le modèle GenSLM de NVIDIA[23]**

Les protéines — s'appuyant sur le succès d'AlphaFold dans la prédiction des structures protéiques, des dizaines d'entreprises affirment maintenant que leurs plateformes d'IA peuvent générer des conceptions de protéines jamais vues auparavant avec de nouvelles caractéristiques commerciales. Certains travaillent même sur des générateurs de texte en protéine.

> **La nouvelle idée est la suivante : puis-je créer un modèle de base qui… parle « protéine » tout comme GPT4 parle anglais ?**
>
> **– Jason Kelly, directeur général de Gingko Bioworks (en partenariat avec Google AI)[24]**

Les éditeurs de gènes — Au cours de la dernière décennie, les outils clés du génie génétique ont été des protéines modifiées telles que CRISPR-Cas 9, que l'on trouvait à l'origine dans des bactéries naturelles. Les entreprises d'IA telles que Profluent utilisent désormais l'IA générative pour générer de nouvelles protéines artificielles issues de l'édition de gènes qui, selon elles, couperont, épisseront et remplaceront l'ADN et l'ARN plus efficacement que CRISPR.[25]

L'épigénétique — Des ingénieurs de l'Université de Toronto, associés à la startup TBG Therapeutics, ont mis au point un système d'IA appelé ZFdesign qui génère des doigts de zinc (ZF). Ceux-ci agissent comme des facteurs de transcription — des éléments épigénétiques qui activent ou désactivent différentes parties de l'ADN. Les chercheurs introduisent des données provenant de milliards d'interactions entre les protéines ZF et l'ADN dans un modèle d'apprentissage automatique, qui génère ensuite des doigts de zinc modifiés qui se lient à une séquence d'ADN donnée.[26] L'IA générative est également utilisée pour la modification des histones[27] — un autre mécanisme épigénétique.

23 Kristen Ye, "The AI Podcast - Anima Anandkumar on Using Generative AI to Tackle Global Challenges – Ep203", NVIDIA, consulté le 25 juillet 2024, https://blogs.nvidia.com/blog/anima-anandkumar-generative-ai/.

24 Jason Kelly s'exprimant sur No Priors Podcast ep34 : "DNA as Code – Cell Programming and AI ", YouTube, téléchargé par No Priors Podcast , https://www.youtube.com/watch?v=snt-fMsCDVI.

25 Callaway E. 2024. "'ChatGPT for CRISPR' creates new gene-editing tools', Nature 629(8011):272. doi: 10.1038/d41586-024-01243-w. PMID: 38684833.

26 Ichikawa DM, Abdin O, Alerasool N, Kogenaru M, Mueller AL, Wen H, Giganti DO, Goldberg GW, Adams S, Spencer JM, Razavi R, Nim S, Zheng H, Gionco C, Clark FT, Strokach A, Hughes TR, Lionnet T, Taipale M, Kim PM, Noyes MB. 2023. "A universal deep-learning model for zinc finger design enables transcription factor reprogramming", Nat Biotechnol 41(8):1117-1129. doi: 10.1038/s41587-022-01624-4. EPUB, 26 janvier 2023. PMID: 36702896; PMCID: PMC10421740.

27 DaSilva LF, Senan S, Patel ZM, Janardhan Reddy A, Gabbita S, Nussbaum Z, Valdez Córdova CM, Wenteler A, Weber N, Tunjic TM, Ahmad Khan T, Li Z, Smith C, Bejan M, Karmel Louis L, Cornejo P, Connell W, Wong ES, Meuleman W, et Pinello L. 2024. "DNA-diffusion: Leveraging generative models for controlling chromatin accessibility and gene expression via synthetic regulatory elements, bioRxiv [Preprint]. 1:2024.02.01.578352. doi: 10.1101/2024.02.01.578352. PMID: 38352499; PMCID: PMC10862870.

Bioproduction - **© Greg Emmerich, Flickr, CC BY-SA 2.0**

2. Bioproduction : stimuler la biotechnologie

Les systèmes d'IA peuvent également redéfinir le matériel, l'infrastructure et le flux de travail de la biotechnologie industrielle — tels que les processus de laboratoire, les installations de fermentation et les cultures cellulaires. Ils peuvent également permettre aux robots de réaliser certaines parties du travail de laboratoire, par exemple en utilisant la vision artificielle à la place des scientifiques humains. La culture de cellules génétiquement modifiées ou la fermentation de bactéries ou de levures modifiées dans un bioréacteur (une grande cuve en acier) est compliquée et sujette à la contamination et à d'autres problèmes. L'utilisation de l'IA pour surveiller et automatiser les processus permettrait aux systèmes de bioproduction de devenir plus efficaces ou d'identifier, de prévenir ou de combattre les contaminants ou les déséquilibres.[28] Si la bioproduction gérée par l'IA peut augmenter la vitesse ou la quantité de production, cela pourrait rendre les produits génétiquement modifiés qui ne pouvaient auparavant pas être produits en quantités au-delà du laboratoire, viables industriellement. Pour ce faire, les gouvernements doivent accroître leur capacité humaine en matière de surveillance, d'évaluation des risques, d'évaluation des technologies, de réglementation, de nettoyage et de rappel. Les mêmes capacités peuvent également être utilisées par des acteurs malveillants pour produire des toxines ou des agents de guerre biologique à double usage en grandes quantités.

28 Amerging Technologies, "How AI is Revolutionizing the Fermentation Industry", 15 mars 2024, https://www.linkedin.com/pulse/how-ai-revolutionizing-fermentation-industry-amergingtech-1dsec/.

3. Systèmes bionumériques : l'exemple des cultures « robotisées »

L'une des caractéristiques du développement technologique actuel est l'essor des systèmes cyberphysiques, c'est-à-dire des systèmes plus complexes du monde réel contrôlés par des technologies numériques, notamment l'IA, les capteurs et l'automatisation. Ces systèmes cyberphysiques incluent de plus en plus d'organismes, de composants et de produits de biologie de synthèse, brouillant la frontière entre les domaines biologique et numérique (connue sous le nom de convergence bionumérique). Un exemple est l'utilisation de cultures génétiquement modifiées et de pulvérisations d'ARNi dans les systèmes d'agriculture numérique gérés par l'IA. Par exemple, la société InnerPlant a bio modifié le soja et d'autres cultures pour émettre une protéine fluorescente lorsqu'ils sont stressés (par exemple, en ayant besoin d'eau ou attaqués par des parasites). Cette fluorescence est réglée pour être détectée par des capteurs montés sur des machines agricoles dotées d'IA, qui peuvent, en réponse, appliquer de l'eau ou des pesticides.[29] En effet, les plantes génétiquement modifiées ont été bioconçues pour être « prêtes pour les robots », c'est-à-dire censées communiquer uniquement avec des machines agricoles exploitées par l'IA. Ensemble, les cultures biomodifiées et les machines agricoles basées sur l'IA permettent à un système cyberphysique d'agriculture numérique plus vaste de changer l'agriculture — un système qui verrouille l'utilisation des pesticides et des mégadonnées.

4. Biocalcul : l'exemple des organoïdes cérébraux

Robots d'IA agricole - © Steve Long, Flickr

Les modèles d'IA sont actuellement basés sur des calculs à grande échelle effectués par des puces de silicium dans des racks de serveurs situés dans de grands entrepôts à température contrôlée, connus sous le nom de centres de données. Ces derniers consomment en permanence d'énormes quantités d'énergie, d'eau et de liquide de refroidissement chimique. Les développeurs de matériel d'IA recherchent activement des solutions de remplacement à faible consommation d'énergie pour les calculs à base de silicium. L'une des orientations futures pourrait consister à exploiter la biologie de synthèse pour fabriquer des cellules vivantes ou de l'ADN qui effectuent des calculs, ce que l'on appelle la bio-informatique. Bien que le rêve de la bio-informatique soit ancien, les coûts massifs en énergie et en ressources de l'IA actuelle basée sur le silicium lui donnent un regain d'intérêt. Dans une démonstration de principe, une société australienne, Cortical Labs, a construit des circuits informatiques d'IA fonctionnels à partir d'organoïdes cérébraux (cellules cérébrales cultivées). Les bio-ingénieurs ont appris à ces circuits d'IA simples et vivants à jouer à des jeux informatiques ou à reconnaître la parole.[30] Si les pièces biologiques synthétiques devenaient courantes en informatique, cela entraînerait des risques de génie génétique et la tâche de gestion de la pollution biologique dans un tout nouveau domaine de l'industrie.

29 "Products", InnerPlant, consulté le 25 juillet 2024, https://innerplant.com/products/./

30 Adhiti Iyer, 13 mars 2024, Is OI the New AI? Questions Surrounding 'Brainoware' Blog - https://blog.petrieflom.law.harvard.edu/2024/03/13/is-oi-the-new-ai-questions-surrounding-brainoware/

La ruée vers la **biologie générative** : les titans de la technologie se lancent dans la biologie (et la politique de la **biotechnologie/biodiversité**)

L'une des caractéristiques les plus frappantes de la biologie de synthèse pilotée par l'IA est qu'une grande partie du travail est dirigée par les plus grandes entreprises technologiques du monde. La plupart de ces nouveaux leaders de l'IA/biologie de synthèse sont des titans du numérique, sans expérience préalable en biotechnologie ou en gestion de la biodiversité, mais une vaste expérience de la mise en œuvre de modèles commerciaux monopolistiques et de contournement des réglementations. Ils concluent des accords conjoints ou acquièrent de petites startups de biotechnologie. Ces entreprises de données agissent à partir de métaphores courantes dans la biologie de synthèse qui prétendent que les systèmes vivants peuvent être manipulés de manière simpliste, comme des machines biologiques qui fonctionnent sur du code et peuvent être programmées et déboguées. Cependant, en réalité, le monde « humide » de la biologie, des écosystèmes et des processus de la vie n'a souvent rien à voir avec l'informatique ou le codage. Ils comportent des risques vivants tout à fait nouveaux. Les titans de la technologie numérique issus d'une culture de la Silicon Valley du « aller vite et tout casser » sont peut-être mal équipés pour gérer les risques biologiques.[31]

Biochimie - **© Day Of Victory Studio, Shutterstock**

31 Le mantra de la Silicon Valley « move fast and break things » est souvent attribué à Mark Zuckerberg de Meta/Facebook — voir https://www.snopes.com/fact-check/move-fast-break-things-facebook-motto/

Participation des géants de la technologie :

Google/Alphabet — Google DeepMind (laboratoire de recherche en IA) a développé le programme de haut niveau AlphaFold. Google a créé une coentreprise avec Gingko Bioworks, une société de biologie de synthèse, pour générer de nouvelles protéines.[32] Ils ont également créé leur propre société de biotechnologie appelée Isomorphic Labs, qui utilise l'IA pour générer de nouveaux composés médicamenteux pour de grandes sociétés pharmaceutiques.[33]

Microsoft — Le directeur général de Microsoft AI, Mustafa Suleyman, a récemment publié un livre très médiatisé (intitulé *The Coming Wave*) sur la manière dont la convergence de l'IA et de la biologie de synthèse va transformer la société (et créer de nouveaux risques).[34] Son entreprise développe plusieurs outils d'IA générative pour la biologie de synthèse, y compris une plateforme médicale générative appelée BioGPT.[35]

Amazon — En juin 2024, le plus grand fournisseur mondial de services cloud de données a annoncé qu'il collaborait avec une société appelée EvolutionaryScale pour héberger ESM3 — une plateforme d'IA de biologie générative entraînée sur « des milliards de séquences de protéines couvrant 3,8 milliards d'années d'évolution ».[36] Selon Amazon, ESM3 peut comprendre des données biologiques complexes provenant de diverses sources et générer des protéines entièrement nouvelles qui n'ont jamais existé dans la nature. Pendant ce temps, le Bezos Earth Fund (associé au fondateur d'Amazon Jeff Bezos et à sa petite amie) a lancé un programme de 100 millions de dollars « l'IA au service du climat et de la nature » qui se concentre sur l'utilisation de l'IA générative pour les protéines alternatives et d'autres matériaux.[37] Amazon s'intéresserait également au biocalcul des organoïdes cérébraux.[38]

NVIDIA — Le plus grand fabricant de puces d'IA au monde est également à l'avant-garde de la biologie générative. Sa plateforme d'IA GenSLM a été entraînée sur des centaines de milliers de génomes pour générer de nouveaux génomes microbiens et viraux. Ils ont notamment développé des séquences candidates de COVID pour les vaccins et semblent avoir prédit avec succès certains nouveaux variants de COVID.[39] NVIDIA collabore également avec la plateforme ESM3 d'Amazon.platform.[40]

32 Jake Wintermute, "Google and Gingko: Foundry-scale Data meets AI" — https://www.ginkgobioworks.com/2023/08/29/google-and-ginkgo-foundry-scale-data-meets-ai/

33 https://www.isomorphiclabs.com

34 Suleyman, M. The coming wave: Technology, power and the 21st century's greatest dilemma. 2023. Crown Publishing Group, Random House, p.55

35 Newton,W. 2023. "What is BioGPT and what does it mean for healthcare?" 9 février,

36 Wood, M. 2024. "Revolutionizing generative biology with AWS and EvolutionaryScale" 25 juin, https://aws.amazon.com/blogs/industries/revolutionizing-generative-biology-with-aws-and-evolutionaryscale/

37 "AI for Climate and Nature", Bezos Earth Fund, consulté le 25 juillet 2024, https://www.bezosearthfund.org/ai-climate-nature

38 Lee, Z. 2023. "This AI startup wants to be the next NVIDIA by building brain cell-powered computers", Forbes, 21 juin, https://www.forbes.com/sites/zinnialee/2023/06/21/cortical-labs-brain-computer/

39 Salian, I. 2023. "Gen AI for the genome: LLM predicts characteristics of COVID variants", Nvidia Blog, 13 novembre, https://blogs.nvidia.com/blog/generative-ai-covid-genome-sequences/

40 Nagel, B. 2024. "NVIDIA and AMS-backed bioresearch startup launch 'ESM3' model'", 26 juin, https://pureai.com/Articles/2024/06/26/Bioresearch-ESM3-Model.aspx

Meta (Facebook) — En 2022, Meta a dévoilé sa plateforme ESMfold en tant que rivale d'AlphaFold de Google. Meta a affirmé que ESMfold abritait plus de 600 millions de structures protéiques et était 60 fois plus rapide qu'AlphaFold.[41] Cependant, le projet a été arrêté en 2023 dans un contexte de licenciements massifs au sein de Meta.[42]

Salesforce — Cette entreprise américaine de premier plan spécialisée dans l'infonuagique a mis au point ProGEN, un LLM d'IA permettant de générer de nouvelles protéines. ProGEN a été formé en introduisant les séquences d'acides aminés de 280 millions de protéines différentes dans un modèle d'apprentissage automatique.[43] En guise de preuve de concept, Salesforce a ensuite ajusté le modèle en l'amorçant avec 56 000 séquences d'une seule classe de protéines : les lysozymes (utilisés pour les ingrédients alimentaires).[44]

Alibaba – En 2023, les scientifiques du géant chinois de la technologie, Alibaba, ont publié les résultats de sa plateforme d'IA LucaProt, qui a été formée pour identifier les virus à ARN. Selon les chercheurs, LucaProt a identifié 161 979 espèces potentielles de virus à ARN et 180 supergroupes de virus à ARN. Ils ont affirmé :

> **Cette étude marque le début d'une nouvelle ère de découverte de virus, fournissant des outils informatiques qui aideront à élargir notre compréhension de la virosphère mondiale de l'ARN et de l'évolution du virus.**[45]

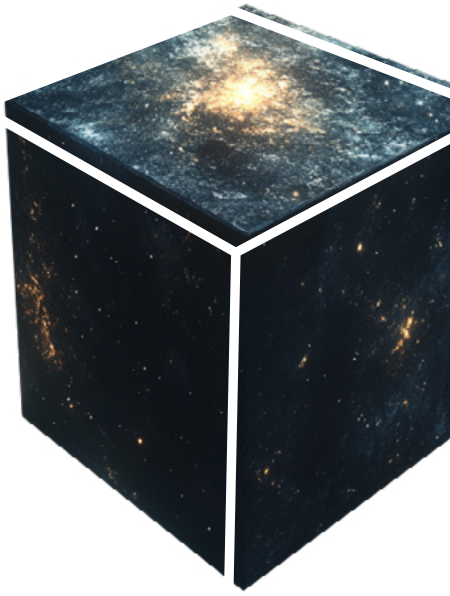
41 Rosenbush, S. 2023. "Scientists at DeepMind, Meta press fusion of AI, biology", *Wall Street Journal* – 22 mars, https://www.wsj.com/articles/scientists-at-deepmind-and-meta-press-fusion-of-ai-biology-4b92af6f

42 Wodecki Jr, B. 2023. "Meta lays off team behind revolutionary protein-folding model', *AI Business*, 8 août.

43 Madani, A. n.d. "ProGen: Using AI to generate proteins", Salesforce AI Research, consulté le 25 juillet 2024, https://blog.salesforceairesearch.com/progen/

44 Université de Californie. 2023. "AI technology generates original proteins from scratch', *Physics.org,* 26 janvier, https://phys.org/news/2023-01-ai-technology-generates-proteins.html

45 Xin Hou, Yong He, Pan Fang, Shi-Qiang Mei et al. 2023. "Artificial intelligence redefines RNA virus discovery", bioRxiv ; doi: https://doi.org/10.1101/2023.04.18.537342



Implications pour la CDB : cinq questions urgentes

Avec l'intérêt croissant des grandes entreprises technologiques, les investissements et le battage médiatique autour de la biologie générative et de l'utilisation de l'IA pour stimuler la production de biotechnologie, les parties à la CDB doivent de toute urgence apporter des réponses aux cinq questions suivantes :

1. La biologie générative de l'IA sape-t-elle les modalités d'accès et de partage des bénéfices du Protocole de Nagoya et la gouvernance des ISN ?

Les dispositions actuelles en matière d'accès et de partage des avantages découlant de l'utilisation des ressources génétiques reposent sur le tracé d'une ligne claire entre l'accès initial à une ressource génétique (par exemple, le prélèvement d'un échantillon de graines ou de tissus) et l'utilisation commerciale finale de ces connaissances (par exemple, dans un produit génétiquement modifié). Cela peut impliquer le traçage de la numérisation d'informations de séquence (par exemple, des codes d'ADN, des séquences de protéines) et leur stockage dans des bases de données d'ISN. Cependant, le fonctionnement de la biologie générative pilotée par l'IA cache tout cela dans une boîte noire algorithmique, brisant tout lien entre les actes d'accès et les actes d'utilisation. De cette façon, on perd l'occasion pour le partage des avantages ou pour les communautés d'exercer un droit de veto sur les utilisations finales.

Au vu des investissements considérables consacrés à la biologie dite générative, il est clair que les géants de la technologie pensent qu'il y a d'énormes profits à tirer de la création de nouvelles protéines, de séquences d'ADN, de virus et d'autres éléments de ce type. Pour générer ces biomolécules, les géants de la technologie s'emparent (gratuitement) de toutes les ISN actuellement disponibles dans les bases de données et les utilisent comme données d'entraînement. Pourtant, lorsqu'il s'agit de partager les avantages, ils prétendent qu'il n'est pas possible d'identifier les ressources génétiques originales spécifiques qui ont amené leurs modèles à suggérer un résultat particulier. En raison de la nature « boîte noire » de l'IA générative, ils affirment que les gouvernements, les régulateurs ou même les experts ne peuvent pas retracer le lien entre l'accès et l'utilisation. Cela entraînera des répercussions sur les négociations actuelles sur les ISN et renforcera l'argument en faveur d'un mécanisme multilatéral et d'un fonds mondial. Les grandes entreprises technologiques qui utilisent des modèles d'IA générative basés sur les ISN devraient être clairement identifiées comme en faisant un usage commercial et devraient cotiser au fonds qui a été créé.

Cette remise en question très sérieuse du paradigme actuel du partage équitable des avantages fait écho à des luttes similaires qui émergent actuellement dans le domaine de la politique en matière d'IA. Des artistes, des écrivains, des journalistes, des photographes et des designers ont intenté des poursuites judiciaires contre des entreprises d'IA qui ont formé leurs modèles sur leurs images et écrits récupérés sur Internet (c'est-à-dire gratuitement).[46] Ils ont même écrit un logiciel pour identifier

46 Un tour d'horizon des affaires de droits d'auteur légaux de l'IA peut être trouvé ici : "Case tracker: Artificial intelligence, copyrights and class actions", BakerHostetler, consulté le 25 juillet 2024, https://www.bakerlaw.com/services/artificial-intelligence-ai/case-tracker-artificial-intelligence-copyrights-and-class-actions/

des œuvres d'art spécifiques dans des ensembles de données d'entraînement.[47] Les entreprises technologiques prétendent qu'elles peuvent échapper aux paiements de droits d'auteur puisque leurs systèmes ne sont « formés » que sur ces intrants et que les résultats qu'ils génèrent sont « nouveaux » et différents. Ils présenteront probablement les mêmes arguments aux titulaires originaux des droits sur les ressources génétiques. Il peut être nécessaire pour les gouvernements d'insister pour qu'une taxe ou une licence soit prélevée sur les entreprises d'IA pour toute utilisation d'ISN dans des ensembles de données d'entraînement. Il peut y avoir des moyens techniques de forcer les entreprises d'IA à maintenir la traçabilité et de forcer la biologie générative à devenir explicable, plutôt que basée sur une boîte noire. Au moins une startup de biologie générative, Basecamp Research basée au Royaume-Uni, affirme déjà que son modèle d'IA est prêt pour Nagoya.[48]

2. La biologie générative de l'IA sape-t-elle les dispositions de biosécurité du Protocole de Cartagena sur la biosécurité ?

Les décisions de la CDB et du Protocole de Cartagena obligent les Parties à entreprendre des évaluations de biosécurité des organismes vivants modifiés et à évaluer les risques des produits, composants et organismes développés par le biais de la biologie de synthèse. Cela est conforme à l'approche de précaution. Pour évaluer l'innocuité biologique d'un nouvel organisme ou d'un nouveau produit, les évaluateurs doivent connaître certains types d'information : ils doivent savoir quelles ressources génétiques ont été ajoutées et de quel organisme elles proviennent (organisme donneur), et comprendre pourquoi ces séquences ont été sélectionnées. Lorsque la bioconception génétique est effectuée dans une boîte noire d'IA, les évaluateurs de la biosécurité peuvent se voir refuser cette information, ainsi que d'autres données importantes, nécessaires à la réalisation d'évaluations de la sécurité et d'autres évaluations (par exemple, socio-économiques).

L'histoire récente de l'IA générative dans d'autres domaines soulève également de sérieux signaux d'alarme indiquant que les organismes ou les protéines conçus par l'IA peuvent ne pas être fiables, qu'ils peuvent comporter des décisions de bioconception inattendues (et inexplicables) et qu'ils sont conçus avec des biais problématiques (voir l'encadré sur les hallucinations, les biais et l'empoisonnement des données de l'IA). La combinaison des risques et des incertitudes associés au génie génétique avec les risques et incertitudes émergents accompagnant l'IA générative crée une recette parfaite pour des catastrophes de type « cygne noir » (risques inhabituels à fort impact qui ne peuvent être prévus).[49]

Une préoccupation plus large pour les régulateurs est ce qu'on a appelé « le problème du rythme ». À mesure que la conception et la construction automatisées d'organismes synthétiques, de virus et de protéines modifiés deviennent plus faciles et moins dépendantes du travail de laboratoire humain, les régulateurs peuvent s'attendre à voir un déluge croissant d'applications pour les OGM et les nouveaux matériaux qui devront être évalués et surveillés. Certains universitaires militaires mettent déjà en garde contre toute initiative visant à confier l'évaluation de la biosécurité des produits et des organismes générés par l'IA à des évaluateurs automatisés de la biosécurité de l'IA — faisant valoir l'importance de garder l'expertise humaine fermement dans la boucle.[50]

47 See https://haveibeentrained.com

48 See https://www.basecamp-research.com/biotech

49 See https://en.wikipedia.org/wiki/Black_swan_theory

50 Vindman, C., et al. n.d. The convergence of AI and synthetic biology: The looming deluge, preprint, consulté le 25 juillet 2024, https://arxiv.org/pdf/2404.18973

Hallucinations, biais et empoisonnement des données de l'IA

Bien que les entreprises d'IA présentent leurs systèmes comme étant capables de générer du texte, des images et des génomes d'organismes vivants, les deux premières années de systèmes d'IA génératifs ont soulevé des problèmes très importants en ce qui concerne la qualité et la précision des résultats. Une partie de ces problèmes découle du fait qu'un système d'IA ne vaut que ce que vaut son ensemble de données d'apprentissage, qui doivent être soigneusement conservées, nettoyées et gérées afin d'obtenir des résultats significatifs et justes. Dans le cas contraire, l'introduction de données erronées engendre la production de données erronées. Les données ne sont pas « pures et immaculées »[51] mais elles peuvent être désordonnées, facilement mal classifiées, manquantes ou fortement biaisées. Plus important encore, les systèmes d'IA générative ne sont pas conçus pour créer ou vérifier la « vérité », mais uniquement pour créer des résultats apparemment crédibles. De cette façon, ils n'atteindront peut-être jamais le niveau de la bonne science.

L'un des résultats problématiques les plus courants de l'IA générative est connu sous le nom d'« hallucinations ». Il s'agit d'un modèle d'IA générative produisant un texte ou une image qui semble avoir un sens, mais qui, après une enquête plus approfondie, est plein d'erreurs, de mensonges et d'éléments improbables. Les images générées par l'IA montreront, par exemple, des personnes avec 6 doigts, du texte brouillé ou des fonctions floues. Le texte généré par l'IA peut rapporter à tort des faits, mal manipuler les chiffres ou inventer des références. En 2023, les analystes ont estimé que les chatbots d'IA hallucinent 27 % du temps, et que près de la moitié (46 %) de leurs réponses contiennent des erreurs factuelles. Si une image ou un texte numérique est rempli de telles erreurs hallucinatoires (comme c'est souvent le cas), cela peut être embêtant ou même drôle, mais si un organisme vivant ou une protéine est créé avec des erreurs hallucinatoires importantes, cela pourrait s'avérer dangereux.

Les chercheurs ont toutefois averti que le terme « hallucinations » pourrait être une métaphore trompeuse, car en fait, l'architecture fondamentale des systèmes d'IA générative signifie que ces systèmes ne sont pas intéressés à produire des résultats véridiques. Les erreurs ne sont pas des erreurs, mais intégrées dans le fonctionnement de l'IA générative. Les chercheurs ont même suggéré que les systèmes d'IA générative devraient être scientifiquement classés comme des « machines à conneries » car ils « ne sont tout simplement pas conçus pour représenter avec précision la façon dont le monde est, mais plutôt pour donner l'impression que c'est ce qu'ils font. »[52]

Une autre source d'erreurs est un biais intégré aux données d'entraînement. Par exemple, il a été constaté que les systèmes de texte d'IA reflètent les préjugés sociaux contre les stéréotypes ethniques et raciaux que l'on trouve dans les données de formation ou qu'ils approfondissent les préjugés sexistes et raciaux dans les décisions d'embauche ou juridiques. Pour cette raison, les entreprises d'IA générative embauchent plusieurs milliers d'annotateurs humains pour « nettoyer » leurs résultats.[53] Dans un contexte de biodiversité, la sélection de données biologiques provenant de certaines cultures commerciales seulement ou provenant d'écosystèmes et de géographies limités pourrait créer des biais problématiques. Il a également été démontré que des entités commerciales ou politiquement motivées peuvent délibérément « empoisonner » des ensembles de données (ajouter ou supprimer des données spécifiques) pour obtenir des résultats biaisés ou pour rendre les options invisibles.[54] L'exercice d'une gouvernance solide sur les ensembles de données d'entraînement devient donc une nécessité politique extrêmement importante mais difficile.

51 Bronson, K. 2022. The immaculate conception of data: Agribusiness, activists, and their shared politics of the future. Montréal: McGill-Queen's University Press

52 Hicks, M.T., Humphries, J. & Slater, J. ChatGPT is bullshit. Ethics Inf Technol 26, 38 (2024). https://doi.org/10.1007/s10676-024-09775-5

53 Dzieza, J. 2023. "AI is a lot of work", *The Verge*, 20 June

54 Lire "Data poisoning: The essential guide", *Nightfall AI*, consulté le 25 juillet 2024, https://www.nightfall.ai/ai-security-101/data-poisoning

3. La biologie générative de l'IA pose-t-elle des risques en matière de biosécurité et d'armes biologiques ?

L'un des risques de la biologie générative de l'IA que certains gouvernements s'efforcent déjà de prendre est la possibilité que ces systèmes génèrent de nouveaux agents pathogènes, toxines ou autres produits qui pourraient être utilisés comme armes. Les modèles de biologie générative de l'IA pourraient être utilisés pour trouver des protéines mortelles ou nuire aux cultures ou pour reconcevoir des virus et des bactéries afin qu'ils se propagent plus facilement ou attaquent certaines espèces. Au moins un article scientifique décrivant un générateur de protéines d'IA décrit comment ce système peut être facilement utilisé pour générer artificiellement des analogues de poisons naturels tels que le venin de serpent.[55]

4. L'intégration de l'IA à la biologie de synthèse améliorera-t-elle ou détériorera-t-elle la santé et la durabilité ?

L'une des raisons pour lesquelles les grandes entreprises d'IA se lancent dans la biologie générative est qu'elles espèrent projeter publiquement la vision d'une gamme potentiellement illimitée de solutions technologiques spéculatives, améliorant ainsi l'acceptabilité sociale de l'IA. La biologie générative pourrait théoriquement être utilisée pour créer de nouvelles alternatives aux protéines animales (Alt-protéines), ce qui, selon ses partisans, pourrait réduire l'empreinte climatique des aliments transformés. On imagine que des protéines d'IA pourraient être développées pour des matériaux qui séquestreront le dioxyde de carbone et que de nouveaux pesticides ou virus conçus par la biologie générative pourraient éradiquer les parasites du paludisme. Les sociétés pharmaceutiques espèrent particulièrement qu'elles pourront utiliser les outils de biologie générative pour créer de nouveaux médicaments ou pour remplacer les composés naturels des produits par de nouveaux composés conçus par la biologie générative qu'elles pourront breveter et vendre.

Pour l'instant, toutes ces promesses sont spéculatives et présentées comme des solutions technologiques trop simplistes. Le bon sens exige que les gouvernements procèdent à des vérifications sur le terrain pour séparer le battage médiatique des promesses, mais aussi à des évaluations technologiques pour comprendre les impacts et les conséquences sociétaux réels et probables. L'histoire de la chimie de synthèse et plus récemment des OGM a montré que les promesses audacieuses de l'industrie doivent faire l'objet d'un examen sociétal et du principe de précaution, au cas où elles s'avéreraient être de fausses solutions, soit ne fonctionnant pas, soit déclenchant des problèmes plus intenses.

Dans le cas de l'intégration de l'IA à la biologie de synthèse, les décideurs politiques doivent être attentifs aux coûts environnementaux réels qui émergent déjà de l'IA générative.[56] Kate Crawford, une éminente chercheuse en IA de l'Université de Californie du Sud, suggère que l'examen de modèles d'IA spécifiques n'aborde pas l'essentiel et que « si nous regardons vraiment cette situation dans son ensemble, il s'agit d'un gigantesque système planétaire ». Pour faire fonctionner l'IA, il faut une

55 Zulfiqar H, Guo Z, Ahmad RM, Ahmed Z, Cai P, Chen X, Zhang Y, Lin H, Shi Z. 2024. "Deep-STP: A deep learning-based approach to predict snake toxin proteins by using word embeddings", Front Med (Lausanne). doi: 10.3389/fmed.2023.1291352. PMID: 38298505; PMCID: PMC10829051

56 Berreby, D. 2024. "As use of AI soars, so does the energy and water it requires", *Yale E360*, 26 février, https://e360.yale.edu/features/artificial-intelligence-climate-energy-emissions

IA Big data - **© frank60, Shutterstock**

énorme quantité de données, ainsi que d'énormes quantités de calcul. Actuellement, ce calcul prend la forme de puces NVIDIA (H100 et A100) fabriquées à Taïwan. Le calcul nécessite de l'énergie, et l'énergie nécessaire pour exécuter ces modèles est stupéfiante — les prédictions sont entre 5 et 50 fois plus énergivores que les fonctions de recherche traditionnelles. La chaleur générée par ce traitement doit être refroidie, ce qui signifie de grandes quantités d'eau. Des chercheurs de l'Université de Californie à Riverside et de l'Université du Texas à Arlington ont estimé qu'environ 700 000 litres d'eau auraient pu être utilisés pour refroidir les machines qui ont formé ChatGPT-3 dans les installations de données de Microsoft (Li et al. 2023). ChatGPT-4 est encore plus énergivore, et Microsoft et Google signalent tous deux que leur consommation d'énergie a augmenté ces dernières années — en hausse de 50 % en 5 ans dans le cas de Google (Malmö, 2024) — en grande partie grâce à l'IA générative.

De plus, le calcul de l'IA est effectué sur des milliers de serveurs électroniques, dont la fabrication nécessite une utilisation intensive de l'eau (pour les puces), des composants de batteries industrielles et l'extraction de métaux, de terres rares et d'autres minéraux de la Terre. Le récent boom de l'IA générative, qui a été accéléré par la sortie de ChatGPT, a déclenché un boom mondial de la construction de centres de données d'IA énergivores. Ceci, à son tour, annule les gains de réduction des émissions de carbone durement acquis dans le secteur de l'énergie et détourne l'eau douce des terres agricoles et des communautés les plus pauvres pour refroidir les puces d'IA. Les plus grandes entreprises d'IA, telles que Microsoft, sont maintenant à l'avant-garde d'une renaissance de la construction d'installations nucléaires, afin de maintenir leurs centres de données en activité.[57] Même la tentative de fournir de l'énergie renouvelable aux centres de données entraîne une expansion problématique de l'extraction de minéraux critiques. Les entreprises d'IA affirment que l'IA deviendra plus économe en énergie au fil du temps et qu'elle peut également contribuer aux efforts de lutte contre le changement climatique. Cependant, on craint de plus en plus que l'IA n'exacerbe les menaces écologiques au moment même où les scientifiques avertissent qu'il reste très peu de temps pour inverser les crises mondiales du climat et de la biodiversité. Le calcul et la formation sur l'ISN génomique peuvent être particulièrement énergivores et gourmands en ressources et ne devraient pas être adoptés à la légère par des parties qui luttent déjà pour protéger la biodiversité, l'eau, la sécurité alimentaire et les objectifs de zéro émission nette.

57 Hiller, J. 2023. "Microsoft targets nuclear to power AI operations", Wall Street Journal, 12 décembre, https://www.wsj.com/tech/ai/microsoft-targets-nuclear-to-power-ai-operations-e10ff798

5. Quelles sont les implications de l'intégration de l'IA à la biologie de synthèse pour les connaissances et les pratiques traditionnelles ?

Le deuxième objectif de la CDB concerne l'utilisation durable de la diversité biologique. Des organismes tels que le Groupe de travail sur l'article 8J jouent un rôle essentiel dans la protection des connaissances et des pratiques traditionnelles qui sous-tendent la conservation et l'utilisation durable de la biodiversité. Les entreprises de biologie générative qui cherchent à proposer des alternatives aux produits naturels, y compris des ingrédients médicinaux et alimentaires, peuvent saper directement les chaînes de valeur durables qui soutiennent des millions de petits agriculteurs dans le monde. Les ressources génétiques numériques qui forment les modèles d'IA sont souvent prélevées à l'origine sur les territoires des peuples autochtones et des communautés locales. Ces ressources sont utilisées dans leurs cérémonies, leurs systèmes alimentaires et leurs modes de vie et existent en relations étroites avec les communautés. Traduire ces connaissances et ces relations en ISN pour former les systèmes d'IA des entreprises à la production d'organismes et de protéines synthétiques commerciaux prolonge les siècles de vol et d'appropriation coloniaux dont les peuples autochtones et les communautés locales ont longtemps souffert. Les chercheurs ont qualifié cette nouvelle vague d'appropriation numérique de la vie de « colonialisme des données ».[58]

Dans les années à venir, les débats sur la biologie générative seront probablement rejointes par d'autres discussions urgentes sur les implications du déploiement de systèmes d'IA dans la foresterie, l'utilisation des terres, la conservation, la gestion marine, la restauration, la gestion des espèces, l'agriculture, etc. Les peuples autochtones et les communautés locales — en tant que détenteurs originaux des ressources biologiques et autres qui sont exploitées, formées et manipulées par l'IA — devraient prendre la tête de la prise de décision politique mondiale autour de l'IA à la CDB et au-delà.

Échange de semences - **© P.Kimeli_CCAFS, Flickr**

58 Couldry, N. & Mejias, U.A. 2020. "The costs of connection: How data are colonizing human life and appropriating it for capitalism", Social Forces, Volume 99, Issue 1, Page e6, https://doi.org/10.1093/sf/soz172

Une piste à suivre pour la CDB

Dans les années à venir, on peut s'attendre à ce que les questions de savoir comment l'IA transformera la conservation, l'utilisation durable et le partage équitable des avantages deviennent un sujet majeur de débat politique et d'action dominant tous les programmes de la Convention. Pour l'instant, le sujet de la biologie générative et de l'intégration de l'IA à la biologie de synthèse est d'une importance urgente et croissante pour les parties.

Le transfert du génie génétique et de la production de nouvelles biomolécules dans la boîte noire de l'IA générative pourrait, en fin de compte, ne pas tenir les promesses scintillantes actuellement faites par les grandes entreprises technologiques. Cependant, dans l'intervalle, les Parties doivent reconnaître que cette approche sape les trois piliers clés qui ont été érigés pour assurer un développement juste et prudent de la biotechnologie : la biosécurité, le partage juste et équitable des avantages, et les moyens de responsabilité et de réparation. Cela soulève également de sérieuses considérations socio-économiques.

Culture du haricot - © Alliance of Bioversity International and CIAT, Flickr

Dans le même temps, l'application d'outils d'IA générative promet d'augmenter le volume et la complexité des organismes et des produits issus de la bio-ingénierie à des niveaux qui pourraient dépasser les capacités réglementaires. Cela arrive exactement au moment où les régulateurs ont le plus besoin de s'assurer qu'ils ont un « humain dans la boucle ». Le battage médiatique et les promesses des géants du numérique qui prennent les commandes de la bioéconomie devront être étudiés et triés — par des êtres humains, et non par des algorithmes — en s'appuyant en particulier sur les connaissances situées des peuples autochtones et des communautés locales. La CDB devra maintenir son engagement historique envers le principe de précaution face à l'approche privilégiée par la Silicon Valley, qui consiste à « aller vite et tout casser » (parfois appelée « principe d'innovation »).[59]

- Lors de la COP 16 de la CDB, les parties ont une occasion très claire et directe d'empêcher la question de la biologie générative de bouleverser et de vider de leur substance les décennies de bon travail que la CDB a mené pour aider à établir une bonne gouvernance en matière de biotechnologie. En suivant les recommandations du mAHTEG sur la biologie synthétique, les parties devraient :
- lancer un processus de formulation de politiques portant sur l'intégration de l'IA à la biologie de synthèse ;
- demander au mAHTEG d'entreprendre une évaluation plus approfondie de l'intégration de l'IA à la biologie de synthèse, conduisant à un rapport abordant, entre autres, les impacts potentiels sur la biosécurité ; l'utilisation durable de la biodiversité ; l'accès équitable et le partage des avantages ; aspects sociaux, économiques et culturels ; les impacts sur les connaissances et les pratiques traditionnelles ; et autres questions pertinentes ;
- demander au Secrétariat de la CDB d'élaborer une série de publications techniques traitant de l'IA et de participer aux activités du système des Nations Unies concernant la gouvernance de l'IA ;
- exhorter les parties à envisager l'élaboration de mécanismes de gouvernance efficaces et équitables pour les ensembles de données d'IA ; les modèles de base ; les outils de bioconception algorithmique ; les outils scientifiques automatisés ; et l'utilisation d'organismes, de composants et de produits de biologie de synthèse dans des systèmes cyberphysiques ;
- insister sur le fait que les grandes entreprises technologiques et les entreprises de biologie générative qui entraînent leurs modèles d'IA sur l'ISN utilisent l'ISN à des fins commerciales, et devraient donc être soumises au mécanisme multilatéral et à l'obligation de contribuer au fonds mondial pour l'ISN.

59 Corporate Europe Observatory. 2018. "The 'innovation principle' trap: Industries behind risky products push for backdoor to bypass EU safety rules", 5 décembre, https://corporateeurope.org/en/environment/2018/12/innovation-principle-trap

Glossaire des termes clés

Accès et partage des avantages — Fait référence à la manière dont les ressources génétiques peuvent être accessibles de manière appropriée avec consentement, et à la manière dont les avantages qui résultent de leur utilisation sont partagés de manière juste et équitable avec les personnes ou les pays qui les fournissent (fournisseurs).

Intelligence artificielle (IA) — Ensemble de technologies basées principalement sur l'apprentissage automatique et l'apprentissage profond. Elles sont utilisées pour l'analyse de données, les prédictions et les prévisions, la catégorisation d'objets, le traitement du langage naturel, les recommandations, l'extraction de données, etc.

Biocalcul — Systèmes de molécules dérivées biologiquement ou de cellules vivantes modifiées qui exécutent des processus de calcul à la place du calcul à base de silicium.

Bioconception — L'application de méthodes et d'approches de conception et d'ingénierie créatives à la conception délibérée de matériaux et de systèmes vivants.

Bioproduction — La croissance, le développement et la récolte de composés industriels (par exemple, de petites molécules ou des protéines) dans des installations biotechnologiques telles que des cuves de fermentation ou des cultures cellulaires. Aussi parfois appelé « biotraitement ».

Biotechnologie — La CDB de l'ONU définit la biotechnologie comme étant « toute application technologique qui utilise des systèmes biologiques, des organismes vivants ou des dérivés de ceux-ci pour fabriquer ou modifier des produits ou des procédés à des fins spécifiques ». Elle fait généralement référence aux biotechnologies moléculaires « modernes » telles que la manipulation et le transfert de gènes, l'édition de gènes, le typage de l'ADN et le clonage de plantes et d'animaux.

Problème de la « boîte noire » — Le manque intrinsèque de transparence ou d'interprétabilité externe des systèmes d'IA dont les modèles internes sont cachés pour être compréhensibles.

Systèmes cyberphysiques — Systèmes plus complexes du monde réel dans lesquels les activités ou les objets du monde réel sont contrôlés par des technologies numériques, y compris l'IA, les capteurs et l'automatisation.

Apprentissage profond — Une forme d'apprentissage automatique qui utilise des réseaux neuronaux multicouches pour traiter les données d'une manière inspirée du cerveau humain.

Modèle de diffusion — Un modèle d'IA générative qui apprend en ajoutant du bruit aléatoire pour détruire les données, puis les reconstruire. Utilisé par exemple pour Stable Diffusion (créateur d'images IA génératives).

Information de séquençage numérique (ISN) sur les ressources génétiques — Terme politique qui fait référence de manière générale aux données de séquençage génétique et à d'autres données numériques connexes. Cela inclut les détails de l'ADN, de l'ARN, des séquences de protéines et d'autres codes biologiques moléculaires d'un organisme qui peuvent être exprimés sous forme numérique.

IA discriminante — Se concentre sur la classification des données existantes. Pour ce faire, il apprend la limite de décision qui sépare les différentes classes de données — par exemple, la classification d'images ou l'identification de la taxonomie des espèces.

IA explicable — Propositions de développement de modèles d'IA qui ne cachent pas leur fonctionnement dans une « boîte noire » — mais stockent plutôt leur fonctionnement là où il est possible de retracer la raison et la source d'un résultat particulier (encore théorique à ce stade).

Modèles de base — Modèles d'IA à grande échelle, larges et adaptables avec de nombreux paramètres, entraînés sur une large gamme d'ensembles de données de manière autosupervisée. Peut être entraîné davantage sur des ensembles de données plus spécifiques pour différentes applications.

IA générative — Modèles d'apprentissage profond qui peuvent prendre des données brutes et générer des sorties statistiquement probables lorsqu'on y est invité. Les modèles génératifs encodent une représentation simplifiée de leurs données d'entraînement et s'en inspirent pour créer une nouvelle œuvre similaire, mais pas identique, aux données d'origine.

Biologie générative — Terme industriel émergent utilisé par les entreprises qui appliquent des modèles d'IA générative entraînés sur des informations de séquence génétique ou biologique pour prédire la conception de nouvelles séquences biologiques (par exemple, des séquences d'ADN ou de protéines).

Hallucination — Une réponse générée par l'IA qui contient des informations fausses ou trompeuses présentées comme des faits. Il peut s'agir d'images ou de sections de texte qui semblent crédibles, mais qui sont impossibles ou erronées.

« L'humain dans la boucle » — Un concept émergent dans le discours de l'IA décrivant les processus qui s'éloignent de l'automatisation complète des machines en veillant à ce que l'action humaine reste au-dessus de la prise de décision et des résultats.

Grands modèles de langage (LLM) — Modèles d'apprentissage automatique qui ingèrent de grandes quantités de texte, à partir desquels ils analysent, classifient, puis génèrent des textes basés sur le langage d'apparence crédible.

Responsabilité et réparation — Système juridique établi en vertu de la CDB, et en particulier du Protocole additionnel de Nagoya-Kuala Lumpur sur la responsabilité et la réparation, qui établit un régime de responsabilité pour les dommages causés à la biodiversité (par exemple par des organismes vivants modifiés) ainsi que des procédures pour assurer le remboursement ou la réparation des dommages.

Apprentissage automatique — L'utilisation de systèmes informatiques capables d'apprendre et de s'adapter sans suivre d'instructions explicites. Les systèmes d'apprentissage automatique utilisent des algorithmes et des modèles statistiques pour analyser et tirer des conclusions à partir de modèles de données.

Biotechnologie moderne — La CDB des Nations Unies définit la « biotechnologie moderne » comme étant « l'application de

- techniques in vitro d'acide nucléique, y compris l'acide désoxyribonucléique (ADN) recombinant et l'injection directe d'acide nucléique dans des cellules ou des organites, ou
- fusion de cellules au-delà de la famille taxonomique, qui surmontent les barrières physiologiques naturelles de reproduction ou de recombinaison et qui ne sont pas des techniques utilisées dans la sélection et la reproduction traditionnelles ».

Le principe ou l'approche de précaution — Une approche en droit de l'environnement qui rééquilibre le fardeau de la preuve dans le cas du risque. La Déclaration de Rio de 1992 définit l'approche de précaution comme suit : « Lorsqu'il existe des menaces de dommages graves ou irréversibles, l'absence de certitude scientifique absolue ne doit pas servir de prétexte pour reporter l'adoption de mesures rentables visant à prévenir la dégradation de l'environnement ».

Modèles de langage protéique à grande échelle — Prot-LLM — LLM entraînés sur des données de séquence liées aux protéines — y compris les séquences d'acides aminés, les modèles de repliement des protéines et d'autres informations biologiques — qui tentent de prédire la structure, les fonctions et les interactions des protéines.

Cultures prêtes à être robotisées — Cultures génétiquement modifiées pour communiquer des informations à l'équipement agricole robotisé (p. ex., elles deviennent fluorescentes lorsqu'elles sont stressées par une attaque d'insectes ou une sécheresse). Les cultures sont conçues pour soutenir l'agriculture numérique.

Biologie de synthèse — définie par la CDB des Nations Unies comme étant « un développement ultérieur et une nouvelle dimension de la biotechnologie moderne qui combine la science, la technologie et l'ingénierie pour faciliter et accélérer la compréhension, la conception, la reconception, la fabrication ou la modification du matériel génétique, des organismes vivants et des systèmes biologiques ».

Données d'entraînement — Ensemble initial de données utilisé pour entraîner un algorithme ou un modèle d'apprentissage automatique.

Modèle basé sur un transformateur — Un exemple de modèle d'IA génératif basé sur un encodeur et un décodeur pour analyser la relation des mots (ou similaires) avec d'autres mots dans une section de texte. Utilisé, par exemple, pour le chatbot d'IA générative ChatGPT.