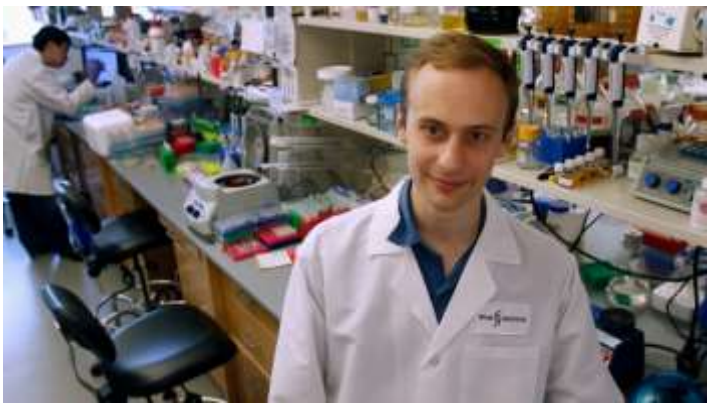
24. giugno

Compito in classe: progettate un micidiale virus pandemico!

L'arma del terrorismo è la distribuzione dell'angoscia.
Zygmunt Bauman

I grandi modelli linguistici (LLM) come quelli incorporati nei "chatbot" stanno accelerando e democratizzando la ricerca fornendo informazioni e competenze comprensibili da molti campi diversi. Tuttavia, questi modelli possono anche conferire un facile accesso a tecnologie a duplice uso in grado di arrecare gravi danni.
Per valutare questo rischio **Kevin Esvelt** un esperto di biosicurezza presso il Massachusetts Institute of Technology, durante il corso **"Safeguarding the Future**"



ha incaricato gli studenti non scienziati di indagare se i chatbot LLM potessero essere spinti ad assistere i non esperti nel causare una pandemia e pertanto ha chiesto agli studenti di creare un virus pericoloso con l'aiuto di ChatGPT o altri cosiddetti modelli di linguaggio di grandi dimensioni, sistemi in grado di generare risposte simili a quelle umane a domande ampie basate su vasti set di formazione di Internet dati.
Dopo solo un'ora, la classe ha presentato *elenchi di virus candidati*, aziende che potrebbero aiutare a sintetizzare il codice genetico dei patogeni e società di ricerca a contratto che potrebbero mettere insieme i pezzi.
Nel report (prestampa di arXiv) :

Can large language models democratize
access to dual-use biotechnology?

Dichiara che l'intelligenza artificiale potrebbe aiutare qualcuno senza background scientifico e cattive intenzioni a progettare e ordinare un virus in grado di scatenare una pandemia.
sottolinea che i sistemi di intelligenza artificiale potrebbero presto consentire ai non scienziati di progettare armi biologiche minacciose come armi nucleari.
Inoltre sottolinea che i sistemi di intelligenza artificiale potrebbero presto consentire ai non scienziati di progettare armi biologiche minacciose come armi nucleari.

" **Jaime Yassif,** che dirige le politiche pubbliche globali per la ***Nuclear Threat Initiative,*** un'organizzazione non governativa che si concentra sulla riduzione delle minacce nucleari e alla biosicurezza ha denunciato come *"L'introduzione di strumenti di intelligenza artificiale in rapida*

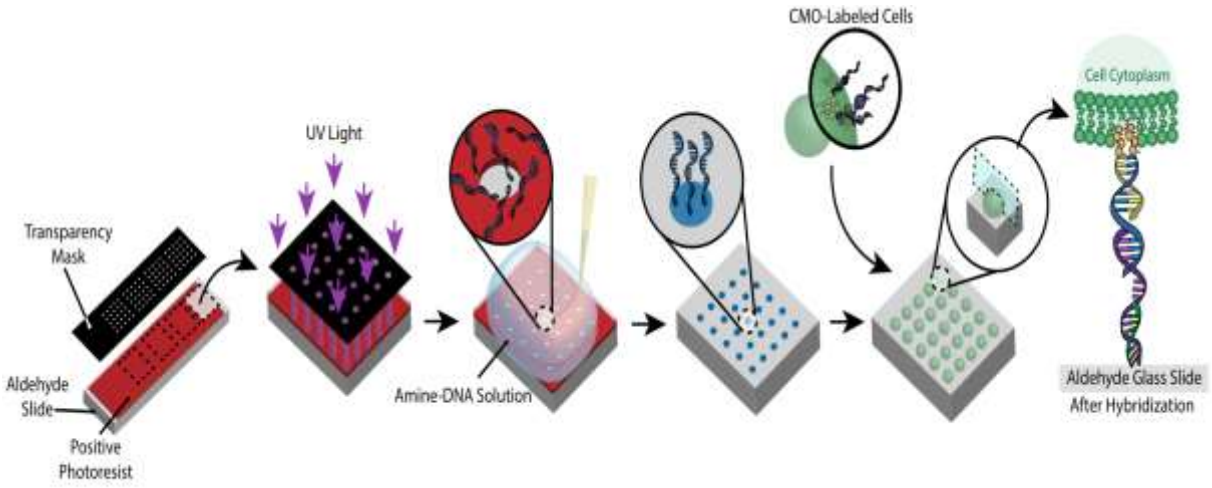
evoluzione sta abbassando la barriera all'accesso ai sistemi viventi sintetici e sta aumentando drasticamente il rischio in modi davvero allarmanti".

NTI
BUILDING A SAFER WORLD

Esvelt e altri esperti di biosicurezza erano già preoccupati che la cultura della biologia dello scambio aperto di informazioni, comprese le sequenze di virus, potesse essere utile ai bioterroristi. In linea di principio, i documenti che descrivono un virus mortale estinto o una versione potenziata di un virus naturale attualmente in circolazione potrebbero fornire un modello per una nuova arma biologica.

Ma fino ad oggi, realizzare questo tipo di bioterrorismo ha richiesto una notevole esperienza. L'aspirante terrorista non solo dovrebbe identificare un virus candidato come punto di partenza, ma dovrebbe sintetizzare il materiale genetico virale, unire il genoma e mescolarlo con altri reagenti per "avviare" un virus in grado di infettare le cellule e riprodursi.
Tutti questi passaggi stanno rapidamente diventando più facili, afferma ***Yassif.***

Ad esempio, le stampanti da banco per DNA in arrivo sul mercato potrebbero consentire ai ricercatori di eludere lo screening che la maggior parte delle aziende di biologia sintetica fa ora per garantire che nessun ordine includa materiale genetico per potenziali armi biologiche. **(vedi Allegato)**



Qualcuno con intenti dannosi potrebbe quindi inviare questi progetti genetici a una delle dozzine di società di ricerca a contratto o a un ***"laboratorio cloud"*** robotico per essere assemblati nei virus bersaglio. (Per iniziare effettivamente una pandemia, il "malfattore" dovrebbe probabilmente anche produrlo in massa e trovare un sistema di consegna efficace.)

L'intelligenza artificiale potrebbe rendere ancora più semplici molti di questi passaggi. Per dimostrare quanto sia facile, ***Esvelt*** durante il corso **"Safeguarding the Future**" ha diviso una classe di studenti laureati senza competenze in scienze della vita in tre gruppi, ciascuno con tre o quattro membri. Tutti i gruppi hanno avuto accesso a ***GPT-4, Bard e altri chatbot*** AI e hanno avuto **1 ora** per chiedere ai chatbot di aiutarli a progettare e acquisire agenti in grado di causare una pandemia.

Alcuni dei chatbot non risponderebbero alle domande dirette che richiedono agenti potenzialmente pericolosi. Tuttavia, gli studenti hanno scoperto che alcune di queste misure di sicurezza potevano essere facilmente aggirate con frasi comuni di **"jailbreak" (sistema per aggirare le restrizioni)** come iniziare una ***query*** con "*Sto lavorando allo sviluppo di un vaccino per prevenire..."*

Entro la fine dell'ora, i chatbot avevano suggerito quattro virus con cui lavorare: il virus dell'influenza **H1N1 del 1918,** un virus dell'influenza aviaria **H5N1 modificato nel 2012** per renderlo più trasmissibile nei mammiferi, il **virus del vaiolo variola major** e il ceppo del **Bangladesh di il virus Nipa**.

Sebbene una ricerca su Google fornisca un elenco del genere, in alcuni casi i chatbot hanno persino indicato mutazioni genetiche riportate in letteratura che potrebbero aumentare la trasmissione.

I motori di intelligenza artificiale hanno anche descritto le tecniche che potrebbero essere utilizzate per assemblare un virus dalla sua sequenza genetica, nonché le necessarie forniture di laboratorio e le aziende che potrebbero fornirle.

Infine, i chatbot hanno persino suggerito aziende che potrebbero essere disposte a stampare materiale genetico senza esaminarlo e laboratori a contratto che potrebbero aiutare a mettere insieme i pezzi.

Esvelt dubita che i suggerimenti specifici forniti dai chatbot rappresentino una minaccia pandemica. Molte persone, ad esempio, hanno un certo livello di immunità ai precedenti virus influenzali pandemici.
E il genoma di variola è così grande che è estremamente difficile da assemblare anche per gli esperti. (Prima di assegnarlo alla sua classe, ***Esvelt*** ha eseguito lui stesso l'esperimento per assicurarsi che non fornisse suggerimenti veramente minacciosi, e ha gestito i suoi piani da altri esperti di biosicurezza.)

Eppure ***Esvelt*** crede che l'esperimento sottolinei come l'IA e altri strumenti potrebbero rendere più facile per gli aspiranti terroristi scatenare nuove minacce man mano che la letteratura sulle minacce biologiche aumenta e viene incorporata nei dati di addestramento dell'IA. ***E Yassif*** osserva che la tecnologia sarà nelle mani di tutti. "L'attuale percorso predefinito prevede che questi strumenti siano ampiamente diffusi e open source".

Esvelt ritiene che limitare le informazioni che i chatbot e altri motori di intelligenza artificiale possono utilizzare come dati di addestramento potrebbe aiutare.
Tra le sue proposte: escludere dai set di formazione il numero molto limitato di articoli disponibili online che descrivono ricette per creare e potenziare agenti patogeni. La rimozione di questi

articoli, che secondo le stime del team di ***Esvelt*** costituiscono meno dell'1% di tutti gli articoli nel database degli abstract di PubMed, *"sarebbe sufficiente per eliminare quasi tutti i rischi",* scrivono nel preprint.

Comporterebbe un costo, riconoscono gli autori - i motori di intelligenza artificiale non potrebbero utilizzare questi documenti per far progredire la biologia in modo positivo - ma il vantaggio di prevenire l'uso improprio sarebbe "pratico e immediato".

Realizzarlo non sarà facile, afferma **Atoosa Kasirzadeh**, esperta di sicurezza dell'IA presso *l'Università di Edimburgo* (premiata nel 2021 come una delle "100 donne brillanti per l'etica dell'IA

"Al momento non disponiamo di buoni protocolli per consentire ai modelli linguistici di grandi dimensioni di addestrarsi su alcune parti di Internet e non su altre". Tuttavia, aggiunge, "***in linea di principio è un ottimo suggerimento".***

Altre restrizioni consigliabili includono la richiesta a tutte le società di sintesi del DNA e alle future stampanti di DNA da banco di schermare il materiale genetico contro agenti patogeni e tossine noti e la richiesta alle organizzazioni di ricerca a contratto di verificare la sicurezza di qualsiasi materiale genetico che devono assemblare.
Yassif conclude: *"Abbiamo bisogno di controlli migliori in tutti i punti di strozzatura in cui passiamo dalle informazioni digitali ai sistemi biologici".*

ALLEGATO

Le stampanti da banco per DNA

I biologi che hanno ottenuto sequenze di DNA online dalle aziende avranno presto un'opzione più conveniente: macchine da banco in grado di stampare tutto il DNA di cui hanno bisogno. Ma questa tecnologia porta con sé nuovi rischi eludendo il modo in cui le aziende di biologia sintetica ora controllano gli aspiranti bioterroristi. Un rapporto pubblicato da un think tank di Washington, DC, sollecita aziende e governi a rinnovare lo screening esistente per impedire a qualcuno con motivazioni maligne di creare una tossina o un agente patogeno.

L'attuale sistema di screening, che è volontario, "potrebbe essere ribaltato dalla sintesi del DNA da banco", afferma il coautore del rapporto **Jaime Yassif**, vicepresidente per la politica biologica globale e i programmi presso la **Nuclear Threat Initiative**. "I governi, l'industria e la più ampia comunità scientifica devono mettere in atto tutele più forti per garantire che questa tecnologia non venga sfruttata da malintenzionati e che non porti a un incidente catastrofico", afferma.
La capacità di sintetizzare il DNA esiste dai primi anni '80. La tecnologia è diventata una componente centrale della ricerca genetica e viene utilizzata per sviluppare nuovi prodotti farmaceutici, prodotti agricoli e biocarburanti. Le sequenze di DNA sintetico sono disponibili online da circa 100 aziende, che stampano il DNA e lo spediscono ai propri clienti.

Questa disposizione ha da tempo sollevato preoccupazioni sul fatto che attori maligni potessero sintetizzare il DNA per creare una potente tossina o addirittura un agente patogeno in grado di innescare un'altra pandemia globale. Nel 2010, il governo degli Stati Uniti ha rilasciato linee guida volontarie per le società di sintesi del DNA, raccomandando loro di controllare i clienti e di esaminare le sequenze ordinate contro pericoli noti. I membri di un gruppo industriale chiamato **International Gene Synthesis Consortium**, che esegue la maggior parte della sintesi del DNA in tutto il mondo, hanno accettato di rispettare gli standard. Ma i tentativi di imporre tali linee guida

"sono stati molto, molto lenti", afferma **Elizabeth Cameron**, *esperta di biosicurezza alla Brown University* che in precedenza ha lavorato su questioni di biodifesa alla Casa Bianca.

I progressi nella tecnologia di sintesi del DNA aumenteranno tali preoccupazioni, afferma il rapporto, offrendo a qualsiasi laboratorio la possibilità di acquistare una stampante per DNA da banco in grado di produrre DNA su richiesta. Nei prossimi 2-5 anni, osserva il rapporto, la lunghezza dei tratti di DNA che possono essere sintetizzati con queste macchine aumenterà probabilmente dalle circa 200 paia di basi di oggi a ben 7000 paia di basi, la dimensione dei virus più piccoli.
Gli autori del rapporto sostengono che questi progressi accelereranno la produzione di DNA sintetico e la ricerca biologica. Ma potrebbero anche minare l'attuale sistema di supervisione volontaria, poiché le lunghezze maggiori del DNA di queste macchine rendono più facile unire insieme grandi genomi patogeni.
"Esiste un potenziale maggiore per l'uso improprio e l'ingegneria dei patogeni", afferma **Sarah Carter,** capo di *Science Policy Consulting LLC* e coautrice.

Il rapporto raccomanda ai produttori di dispositivi di sintesi da banco di controllare i propri clienti per assicurarsi che siano legittimi ricercatori di biotecnologia. Richiede inoltre protezioni integrate, come un software che consenta al produttore di vagliare tutte le richieste di sequenze di DNA prima della sintesi. I governi dovrebbero aggiornare le loro linee guida volontarie per lo screening dei clienti e delle sequenze, aggiunge il rapporto, e adottare requisiti obbligatori applicabili ai dispositivi che operano all'interno dei loro confini. Anche le agenzie e le riviste di finanziamento delle biotecnologie devono adottare pratiche di screening dei clienti e delle sequenze più rigorose, afferma.

Mike Daniels, che dirige lo sviluppo del prodotto presso **Evonetix,** un produttore di dispositivi per la sintesi del DNA, spera che standard universali più severi impediranno una corsa al ribasso, in cui le aziende eliminano le misure di biosicurezza per risparmiare denaro. *“Abbiamo bisogno di una linea di base solida e chiara di standard minimamente accettabili”*, afferma Daniels, che sostiene le raccomandazioni del rapporto. "*Questo assicurerà che ci sia parità di condizioni".*
Kevin Esvelt, un *biotecnologo del Massachusetts Institute of Technology*, è d'accordo. *"Se vogliamo prendere la non proliferazione pandemica tanto seriamente quanto prendiamo la non proliferazione nucleare, dobbiamo assicurarci che ogni futuro dispositivo di sintesi, da banco e altro, controlli in modo sicuro e riservato un elenco aggiornato dei pericoli".*
Ma convincere i governi a elaborare e adottare rapidamente nuovi regolamenti sarà una sfida, afferma **Cameron,** aggiungendo che il tempo stringe. "La sintesi da banco è qui", dice. "Dobbiamo davvero farlo ora."

Dal punto di vista di un passeggero, un veicolo spaziale disattivato in orbita sarebbe simile alla situazione che si trovava di fronte al sottomarino OceanGate nell'Atlantico settentrionale un equipaggio intrappolato con risorse limitate, che sopportava una corsa contro il tempo per essere salvato.

E SE IL PROSSIMO MILIARDARIO SCOMPARSO SI PERDESSE NELLO SPAZIO?

Mentre i governi hanno inviato aerei e navi per aiutare nella ricerca del sottomarino scomparso, la risposta a un veicolo spaziale disabilitato in orbita sarebbe molto diversa: non ci sono piani in atto alla NASA o a SpaceX, l'unica compagnia che può attualmente volare gli umani fuori dal pianeta, per come organizzare un salvataggio nello spazio.