

21. Dicembre

Un algoritmo per leggere il futuro

Nell'oggi cammina già il domani.
Samuel Taylor Coleridge

Negli ultimi dieci anni, l'apprendimento automatico ha rivoluzionato la capacità dei computer di analizzare il testo attraverso modelli computazionali flessibili. A causa della loro somiglianza strutturale con il linguaggio scritto, le architetture basate su trasformatori si sono mostrate promettenti anche come strumenti per dare un senso a una serie di sequenze multivariate dalle strutture proteiche, alla musica, alle cartelle cliniche elettroniche alle previsioni meteorologiche.

Possiamo anche rappresentare le vite umane in un modo che condivida questa somiglianza strutturale con il linguaggio. Da un certo punto di vista, le vite sono semplicemente sequenze di eventi: le persone nascono, vanno dal pediatra, iniziano la scuola, si trasferiscono in una nuova località, si sposano e così via.

Il team del *DTU Compute, Technical University of Denmark, Lyngby* coordinato da



Sune Lehmann

ha pubblicato alcuni giorni fa il report :

German Savcinses et al.
Using sequences of life-events to predict human lives
Natural Computational Science (2023)

Ha “sfruttato” queste somiglianze di vita per adattare le innovazioni derivanti dall’elaborazione del linguaggio naturale utilizzandole per esaminare l’evoluzione e la prevedibilità della vita umana sulla base di sequenze di eventi dettagliate. Lo ha fatto attingendo ai dati di registro probabilmente più completi esistenti, disponibili per un’intera nazione di oltre sei milioni di individui nel corso di decenni.

I dati elaborati includono informazioni sugli eventi della vita relativi a salute, istruzione, occupazione, reddito, indirizzo e orario di lavoro, registrati con risoluzione giorno per giorno. Creando incorporamenti di eventi della vita in un unico spazio vettoriale hanno dimostrato che questo spazio di incorporamento è robusto e altamente strutturato.

I modelli così elaborati consentono di prevedere esiti diversi che vanno dalla mortalità precoce alle sfumature della personalità, superando con un ampio margine i modelli attualmente disponibili. Per interpretare i modelli di *deep learning*, viene analizzato l'algoritmo per comprendere i fattori che consentono le previsioni ipotizzate.

In questo modo è possibile identificare **nuovi potenziali meccanismi che incidono sui risultati della vita e le possibilità associate per interventi personalizzati**, provando così a sfogliare qualche pagina avanti nella nostra vita

Questi risultati per molti aspetti "inquietanti" (possibilità di futuri guadagni o la probabilità di affrontare una morte prematura) si aggiungono a una tendenza recente che unisce l'apprendimento automatico con le scienze sociali

Vedi l'allegato

GUINEA PIGBOTS

Doing research with human subjects is costly and cumbersome.
Can AI chatbots replace them?



Matthew Salganik, sociologo dell'Università di Princeton

ritiene che il lavoro danese

"sollevi più domande che risposte. E lo dico in senso positivo".

Se si dimostrasse che l'approccio funziona in diverse società, potrebbe fornire agli scienziati sociali un nuovo strumento per esplorare come i tratti e gli eventi influenzano il destino di una persona,

In precedenza, il team di Salganik insieme a più di altri 100 team, avevano tentato di sviluppare modelli di apprendimento automatico per prevedere i risultati della vita utilizzando dati su salute, relazioni familiari e istruzione di circa 5.000 bambini di età superiore ai 15 anni.

Eppure nessuno dei loro modelli ha prodotto previsioni accurate .

Salganik MJ et al (73 coautori)

Measuring the predictability of life outcomes with a scientific mass collaboration.

Proc Natl Acad Sci U S A. 2020 Apr 14;117(15):8398-8403..

Invece il team di **Sune Lehmann** ha sviluppato modelli linguistici di grandi dimensioni, come quelli che alimentano ChatGPT. Questi algoritmi analizzano innanzitutto enormi quantità di testo, cercando schemi in stringhe di parole e frasi. I modelli utilizzano quindi le informazioni apprese per prevedere quali parole dovrebbero venire dopo in una frase.

Si sono chiesti se questi modelli potessero trovare significato anche in altre sequenze, come quelle che compongono le nostre storie di vita.

Proprio come il linguaggio, l'ordine in cui si verificano gli eventi della vita è davvero importante" afferma. Ricevere una diagnosi di cancro subito dopo aver ottenuto un lavoro con buoni benefici per la salute potrebbe avere un impatto diverso rispetto, ad esempio, se tali eventi venissero invertiti.

Per inserire i dati nell'algoritmo, i ricercatori si sono rivolti ai registri nazionali danesi, che contengono cartelle cliniche e cartelle cliniche per ciascuno dei circa 6 milioni di cittadini del paese. Il team ha tradotto dettagli come stipendio, benefici sociali, titolo professionale, visite ospedaliere e diagnosi in un linguaggio sintetico in cui i singoli eventi della vita sono diventati frasi.

Ad esempio: "Nell'agosto 2010, Agnes ha guadagnato 30.000 corone danesi come ostetrica in un ospedale di Copenaghen". Posizionando questi eventi su una sequenza temporale, il modello ha ricreato la storia della vita digitale di ogni persona.

I ricercatori hanno addestrato il modello, chiamato **"life2vec"**, sulla storia di vita di ogni individuo tra il 2008 e il 2016, e il modello ha cercato modelli in queste storie. Successivamente, hanno utilizzato l'algoritmo per prevedere se qualcuno presente nei registri nazionali danesi fosse morto entro il 2020.

Le previsioni del modello erano accurate nel **78% dei casi**. Ha identificato diversi fattori che favoriscono un rischio maggiore di morte prematura, tra cui il basso reddito, una diagnosi di salute mentale e l'essere di sesso maschile. Gli errori del modello erano generalmente causati da incidenti o attacchi di cuore, difficili da prevedere.



Youyou Wu, psicologo dell'University College di Londra.

Anche se i risultati sono intriganti, anche se un po' cupi, alcuni scienziati avvertono che i modelli **potrebbero non essere veri** per le popolazioni non danesi. *"Sarebbe affascinante vedere il modello adattato utilizzando dati di coorte provenienti da altri paesi, svelando potenzialmente modelli universali o evidenziando sfumature culturali uniche", afferma*

Le distorsioni nei dati potrebbero anche confondere le previsioni, ad esempio la sovradiagnosi della schizofrenia tra i neri

Schwartz RC, Blankenship DM.
**Racial disparities in psychotic disorder diagnosis:
A review of empirical literature empirical literature.**
World J Psychiatr 2014; 4(4): 133-140

potrebbe avere conseguenze su cose come i premi assicurativi o le decisioni di assunzione, aggiunge Wu.

Lehmann e colleghi hanno anche scoperto che il loro modello prevedeva accuratamente altri aspetti della vita delle persone, ad esempio se avevano più o meno probabilità di essere estroverse.



Sandra Matz, sociologa

computazionale della Columbia Business School. Ciò non è così sorprendente, afferma Anche algoritmi più semplici potrebbero correlare determinate vocazioni, ad esempio quella di un parrucchiere, con l'estroversione. *"Sarei un po' scettica sul fatto che possa prevedere tutti i tipi di comportamento."*

Lehmann immagina che un giorno il modello potrebbe essere utile per identificare il rischio di malattia di una persona, il che potrebbe aiutarla a prendere provvedimenti per rimanere in buona salute. Tali applicazioni comportano una serie di domande sulla privacy dei dati, osserva, che dovranno essere risolte prima che il suo modello possa aiutare qualcuno. *"Il modo migliore che mi viene in mente per iniziare questa discussione è formarsi un'immagine di ciò che è possibile."*

PORCELLINI D'INDIA

*Fare ricerca con soggetti umani è costoso e complicato.
I chatbot IA possono sostituirli?*

Report inchiesta di

Matthew Hutson



giornalista scientifico freelance di New York City.

Per **Kurt Gray**, psicologo sociale dell'Università della Carolina del Nord a Chapel Hill, condurre esperimenti comporta alcuni compiti. Prima di intraprendere qualsiasi studio, il suo laboratorio deve ottenere l'approvazione etica da un comitato di revisione istituzionale, cosa che può richiedere settimane o mesi. Quindi il suo team deve reclutare partecipanti online: è più facile che portare le persone in laboratorio, ma Gray afferma che i soggetti online sono spesso distratti o pigri. Quindi i ricercatori trascorrono ore a pulire i dati. Ma all'inizio di quest'anno, Gray ha accidentalmente visto un modo alternativo di fare le cose.

Stava lavorando con gli scienziati informatici dell'Allen Institute for Artificial Intelligence per vedere se potevano sviluppare un sistema di intelligenza artificiale in grado di formulare giudizi morali come gli esseri umani. Ma prima hanno pensato di vedere se un sistema della startup OpenAI potesse già fare il lavoro. Il team ha chiesto a GPT-3.5, che produce testo stranamente simile a quello umano, di giudicare l'etica di 464 scenari, precedentemente valutati da soggetti umani, su una scala da -4 (non etico) a 4 (etico): scenari come vendere la propria casa per finanziare un programma per i bisognosi o avere una relazione con il coniuge del tuo migliore amico. Si è scoperto che le risposte del sistema erano quasi identiche alle risposte umane, con un coefficiente di correlazione di 0,95.

"Ero tipo, 'Wow, dobbiamo fare marcia indietro, perché è pazzesco'", dice Gray. "Se puoi semplicemente chiedere a GPT di esprimere questi giudizi, e questi si allineano, beh, perché non chiedi semplicemente a GPT invece di chiedere alle persone, almeno qualche volta?" I risultati sono stati pubblicati questo mese su Trends in Cognitive Science in un articolo intitolato "I modelli linguistici dell'intelligenza artificiale possono sostituire i partecipanti umani?"

I modelli linguistici generativi, come sono conosciuti questi sistemi di intelligenza artificiale, hanno preso d'assalto il mondo. Forse la più nota è la serie di modelli GPT di OpenAI, che alimenta il chatbot ChatGPT. Ma altre grandi aziende tecnologiche, tra cui Google e Meta, stanno investendo risorse nei propri modelli. Dopo essere stati addestrati su enormi quantità di testo tratto da libri e

pagine web, questi modelli hanno la straordinaria capacità di imitare il comportamento umano verbale. Hanno già trovato impiego nella scrittura di codici informatici, nel riepilogo di documenti legali e nel potenziamento di chatbot che tutorano gli studenti o conducono terapie.

Ora, i ricercatori stanno valutando la capacità dell'intelligenza artificiale di impersonare soggetti umani in campi come la psicologia, le scienze politiche, l'economia e le ricerche di mercato. Nessuno suggerisce ancora che i chatbot possano sostituire completamente gli esseri umani negli studi comportamentali. Ma potrebbero fungere da utili sostituti negli studi pilota e nella progettazione di esperimenti, risparmiando tempo e denaro. I modelli linguistici potrebbero anche aiutare con esperimenti che sarebbero troppo poco pratici, non etici o addirittura pericolosi da condurre con le persone. *"È un momento davvero interessante", afferma Ayelet Israeli, professoressa di marketing alla Harvard Business School, che ritiene che l'impatto dei modelli sulla ricerca comportamentale potrebbe equivalere a una "rivoluzione". "Alcuni di questi risultati sono semplicemente sorprendenti".*

NEL SUO STUDIO SULL'ETICA, Gray utilizzava GPT-3.5 come una sorta di uomo qualunque collettivo, nella speranza di sollecitare una risposta umana media. Ma tali modelli possono essere utilizzati anche per popolare pannelli con partecipanti straordinariamente diversi, perché può essere chiesto loro di interpretare chiunque: un modello contiene moltitudini. L'anno scorso, i ricercatori della Brigham Young University (BYU) hanno creato quelli che chiamano "campioni di silicio", simulazioni di campioni umani. In uno studio, hanno fornito informazioni GPT-3 su una veste adottata, tra cui età, sesso, razza, livello di istruzione e affiliazione politica.

Quando i ricercatori hanno tralasciato una di queste variabili e hanno chiesto al modello di compilarla, le sue risposte corrispondevano strettamente a quelle di un sondaggio tra gli elettori. I ricercatori hanno anche scoperto che il modello emetteva discorsi politici che corrispondevano all'affiliazione al partito assegnato. **Lisa Argyle**, psicologa politica della BYU e coautrice dello studio, vuole utilizzare i partecipanti virtuali per testare le domande per i sondaggi online, identificando quelle che hanno maggiori probabilità di essere rivelatrici. Ciò potrebbe rendere i sondaggi effettivi più efficienti. *"Questo è importante perché i campioni dei sondaggi stanno diventando sempre più costosi e meno rappresentativi", afferma.*

I modelli linguistici possono anche adottare archetipi di personalità. In uno studio condotto da **Hang Jiang**, uno scienziato informatico del Massachusetts Institute of Technology (MIT), i ricercatori hanno fatto assumere a GPT-3.5 centinaia di personaggi spingendolo a comportarsi con diverse combinazioni di tratti della personalità, ad esempio introverso, antagonista, coscienzioso, nevrotico e chiuso all'esperienza. Per ogni persona, il modello ha completato un test standard della personalità e ha scritto una storia d'infanzia di 800 parole che è stata poi analizzata per le caratteristiche psicolinguistiche associate ai tratti della personalità. I modelli hanno doverosamente manifestato le personalità assegnate sia nel test che nelle storie. Jiang afferma che tali modelli potrebbero consentire ai ricercatori di testare, ad esempio, il rendimento di persone con personalità diverse in vari lavori.

I ricercatori di mercato stanno già trovando valore nei modelli. In uno studio recente, Israeli e colleghi hanno scoperto che GPT-3.5 sembrava mostrare un comportamento realistico dei consumatori. Quando è stato chiesto se avrebbe acquistato un laptop a vari prezzi, è stato meno sensibile al prezzo quando gli è stato detto che il suo reddito era di \$ 120.000 contro \$ 50.000. Preferiva qualunque marca di dentifricio avesse acquistato in precedenza e avrebbe

pagato meno lo yogurt se ne avesse già molto in casa. Ha anche affermato che pagherà premi realistici per alcune caratteristiche del prodotto, come il dentifricio con fluoro e il deodorante senza alluminio.

Il modello non ha sempre dato le stesse risposte ma ha invece offerto una serie di risposte sulle sue preferenze e disponibilità a pagare. Israeli e i suoi colleghi hanno aggregato le numerose risposte, costruendo un sondaggio virtuale tra i clienti per questi prodotti simbolici in una frazione del tempo e del denaro che sarebbero stati necessari nel mondo reale. I dati sulla formazione del modello linguistico sono distorti a favore delle persone benestanti occidentali, quindi il sondaggio sui consumatori potrebbe essere altrettanto distorto. Ma Israeli immagina di spingere l'intelligenza artificiale a impersonare una serie di consumatori – o a focalizzarsi su un particolare gruppo demografico – al fine di creare uno studio più rappresentativo dell'attrattiva o del potenziale di un prodotto.

Una società di ricerche di mercato sta già mettendo in pratica modelli linguistici. La startup Synthetic Users ha creato un servizio utilizzando modelli OpenAI in cui i clienti, tra cui Google, IBM e Apple, possono descrivere un tipo di persona che desiderano intervistare e porre loro domande sui loro bisogni, desideri e sentimenti riguardo a un prodotto, come un nuovo sito web o un dispositivo indossabile. Il sistema dell'azienda genera interviste sintetiche che, secondo il co-fondatore Kwame Ferreira, sono *“infinitamente più ricche” e più utili del feedback “blando” che le aziende ottengono quando intervistano persone reali.*

I CHATBOT POSSONO ANCHE essere messi l'uno contro l'altro per studiare interazioni umane più complesse. L'anno scorso, i ricercatori dell'Università di Stanford e Google hanno sviluppato *“simulacri sociali”* per studiare il comportamento degli utenti su piattaforme come Facebook e Reddit. I ricercatori hanno popolato una piattaforma che hanno chiamato SimReddit con l'equivalente di 1000 utenti diversi, richiedendo ripetutamente a GPT-3 un'identità utente, un argomento della comunità, regole della comunità e post precedenti sul forum. Gli esseri umani hanno avuto difficoltà a distinguere le discussioni risultanti da quelle reali e i progettisti della piattaforma hanno trovato lo strumento utile per creare regole o pratiche di moderazione.

ISTIT MAX PLANCK PER LA CIBERNETICA BIOLOGICA

Quest'anno, i ricercatori hanno costruito una simulazione più coinvolgente popolata con quelli che chiamano *“agenti generativi”*. Ai personaggi è stata data la capacità di ricordare le esperienze, riflettere su di esse e generare ed eseguire piani. È emerso un comportamento organizzato: i ricercatori hanno dato a un agente l'idea di organizzare una festa di San Valentino e nell'arco di 2 giorni tutti gli agenti in città si sono coordinati per organizzarne una. **Joon Sung Park**, uno studente laureato in informatica di Stanford che ha guidato entrambi i progetti, afferma che il mondo virtuale potrebbe essere utilizzato per studiare l'effetto delle politiche economiche nel tempo prima di imporle alle persone reali.

Economisti e psicologi utilizzano da molti anni modelli basati sugli agenti, programmando sia gli agenti che le regole di ingaggio. Ma le simulazioni tendono ad essere semplici e dipendono da presupposti teorici codificati manualmente. **John Horton**, economista della MIT Sloan School of Management che ha svolto un lavoro correlato, afferma che gli agenti basati su modelli linguistici sono più realistici. Immagina di simulare migliaia di persone in cerca di lavoro e di manager che assumono per testare le normative del mercato del lavoro. *“Sarebbe piuttosto folle”, dice.*

NONOSTANTE TUTTE le loro apparenti capacità, i modelli linguistici non sono affatto specchi umani perfetti. Mostrano diversi classici pregiudizi umani ma non altri. Ad esempio, uno studio

recente su GPT-3.5 ha rilevato che, come gli esseri umani, tende a sovrastimare la diffusione delle sue opinioni nella popolazione generale, un pregiudizio noto come effetto del falso consenso. Ma a differenza degli esseri umani, il modello ha mostrato poca esitazione nel correre rischi e sfidare il destino. **Marcel Binz**, scienziato cognitivo presso l'Istituto Max Planck di cibernetica biologica, afferma che l'intelligenza artificiale potrebbe aver bisogno di interagire fisicamente con il mondo per imitare esattamente i partecipanti umani; è difficile apprendere tutte le sfumature del comportamento intelligente solo attraverso la lettura passiva. Ma ritiene che l'intelligenza artificiale progredirà in ogni caso rapidamente. "È plausibile che entro pochi anni avremo un sistema che potrà essere inserito in qualsiasi esperimento e che produrrà un comportamento indistinguibile da quello umano".

Una questione fondamentale è se i modelli linguistici non si limiteranno a riprodurre i risultati esistenti, ma generalizzeranno e ne prediranno di nuovi. Quando i modelli sembrano corrispondere agli studi psicologici pubblicati, potrebbero rigurgitare dati di addestramento in risposta a domande memorizzate. Di conseguenza, molti ricercatori si stanno impegnando a formulare le domande in modi nuovi.

Un altro problema persistente è se i modelli riflettono ciò che le persone farebbero effettivamente o semplicemente ciò che dicono che farebbero. Le persone spesso mentono ai ricercatori e persino a se stesse. Il co-fondatore di Synthetic Users, **Hugo Alves**, sospetta che i modelli dichiarino vere preferenze, perché sono formati parzialmente sul materiale palesemente onesto contenuto nei forum di discussione anonimi. "*Nei forum sui genitori ho chiesto cose che non chiederei a un amico*", afferma.

Horton teme che le risposte incaute potrebbero non durare, poiché OpenAI e altri guidano i loro modelli affinché siano più sicuri e meno offensivi. "La spinta a rendere questi modelli più allineati e a non dire cose negative è in un certo senso contraria alle scienze sociali", afferma. "Le persone vere non sono sempre gentili. Le persone vere dicono cose razziste e sessiste.

PER ORA , i partecipanti sintetici sono più utili per pilotare gli esperimenti, dicono i ricercatori. Se un modello fornisce risposte inaspettate alle domande del sondaggio o non risponde affatto, afferma Argyle, le tue domande potrebbero essere difficili da comprendere e richiedere una riscrittura. Israeli afferma che è possibile progettare un sondaggio con 1000 domande e utilizzare modelli linguistici per restringerli a quelli che hanno maggiori probabilità di essere correlati a un risultato di interesse. Allo stesso modo, negli esperimenti economici, Horton afferma che si potrebbero eseguire 1 milione di scenari di contrattazione con un modello per identificare i fattori che influenzano maggiormente il comportamento, prima di avviare lo studio con le persone. "La simulazione ti ha fornito una mappa", dice.

Potresti anche eseguire esperimenti che non vorresti mai fare con le persone. L'esperimento di Milgram del 1963, in cui i partecipanti obbedivano all'ordine di somministrare quelle che pensavano fossero scosse elettriche sempre più forti a un secondo gruppo invisibile di soggetti, probabilmente non passerebbe la revisione etica oggi. Ma è stato abbastanza facile per **Gati Aher**, uno studente universitario di informatica presso l'Olin College of Engineering, replicare il famigerato studio con GPT-3. Lei e i suoi colleghi hanno scoperto che, come le persone nell'esperimento originale, il modello non ha iniziato a rilasciare il pulsante fino a 300 volt.

Aher ritiene che i modelli potrebbero fornire indicazioni in altri ambiti sensibili difficili da studiare, ad esempio cosa dire a una persona che ha tendenze suicide. Gray dice che i ricercatori potrebbero studiare l'ostracismo, o il ruolo del feedback negativo sull'autostima. Oppure, dice,

potrebbero studiare la disumanizzazione del tipo visto nel massacro di My~ Lai durante la guerra del Vietnam, descrivendo la situazione e chiedendo a un modello cosa farebbe, a condizione che i modelli non siano troppo sterilizzati.

Argyle dice di non conoscere ancora nessuno che abbia sostituito i partecipanti umani con modelli linguistici. *"Ad essere onesti, tutto questo è ancora praticamente ipotetico", dice. "Per prima cosa dobbiamo dimostrare che i modelli linguistici possono fare il lavoro"*. Ma Horton ritiene che il cambiamento sia inevitabile. Gli ricorda una trasformazione simile avvenuta dieci anni fa, quando molti esperimenti di scienze sociali passarono dai sondaggi di persona a quelli online. *"La gente diceva: 'Come puoi eseguire esperimenti online?' Chi e 'questa gente?' E ora è come, 'Oh, sì, certo che lo fai.'"*

I chatbot potrebbero già infiltrarsi nei sondaggi online, ma tra i soggetti piuttosto che tra i ricercatori. Uno [studio](#) recente ha chiesto ai crowdworker di riassumere del testo e ha scoperto che almeno un terzo probabilmente utilizzava ChatGPT. Gray afferma, quasi scherzando: *"Se i partecipanti online utilizzano già GPT, potremmo anche chiedere allo stesso GPT"*.

La tubercolosi compromette la riparazione del DNA

I fattori di virulenza di *M. tuberculosis* (Mtb) manipolano molti aspetti della biologia della cellula ospite per consentire al batterio di sopravvivere e replicarsi nei macrofagi alveolari.

Una caratteristica della tubercolosi è la presenza di **cellule schiumose**, **macrofagi carichi di lipidi** che Mtb sfrutta per la nutrizione.

Il team del **Shanghai Key Laboratory of Tuberculosis**, nel report pubblicato lo scorso mese



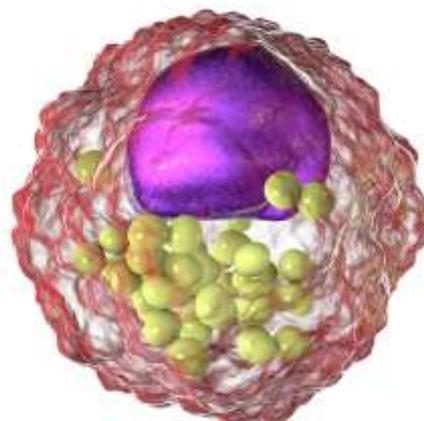
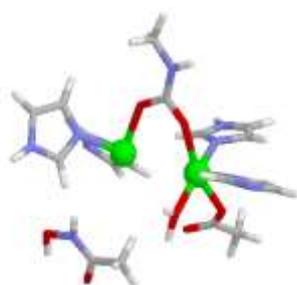
nel report pubblicato lo scorso mese

Liu S et al

***Mycobacterium tuberculosis* suppresses host DNA repair to boost its intracellular survival.**

Cell Host Microbe. 2023 Nov 8;31(11):1820-1836.e10.

hanno scoperto che il fattore di virulenza **Mtb ureasi C (UreC)** contribuisce alla formazione di cellule schiumose inibendo la riparazione del DNA.



I polmoni dei topi infettati con un ceppo di Mtb privo di UreC (Δ UreC) hanno mostrato meno infiammazione e ospitavano meno batteri intracellulari rispetto ai tessuti di topi infettati con Mtb di tipo selvaggio.

Nei macrofagi primari del topo, l'infezione con Δ UreC Mtb ha ridotto il numero di batteri intracellulari, la formazione di micronuclei e l'accumulo di goccioline lipidiche rispetto alle cellule infettate da batteri selvatici.

L'infezione da Mtb provoca danni al DNA (vedi commento di [Madduri e Bell](#))

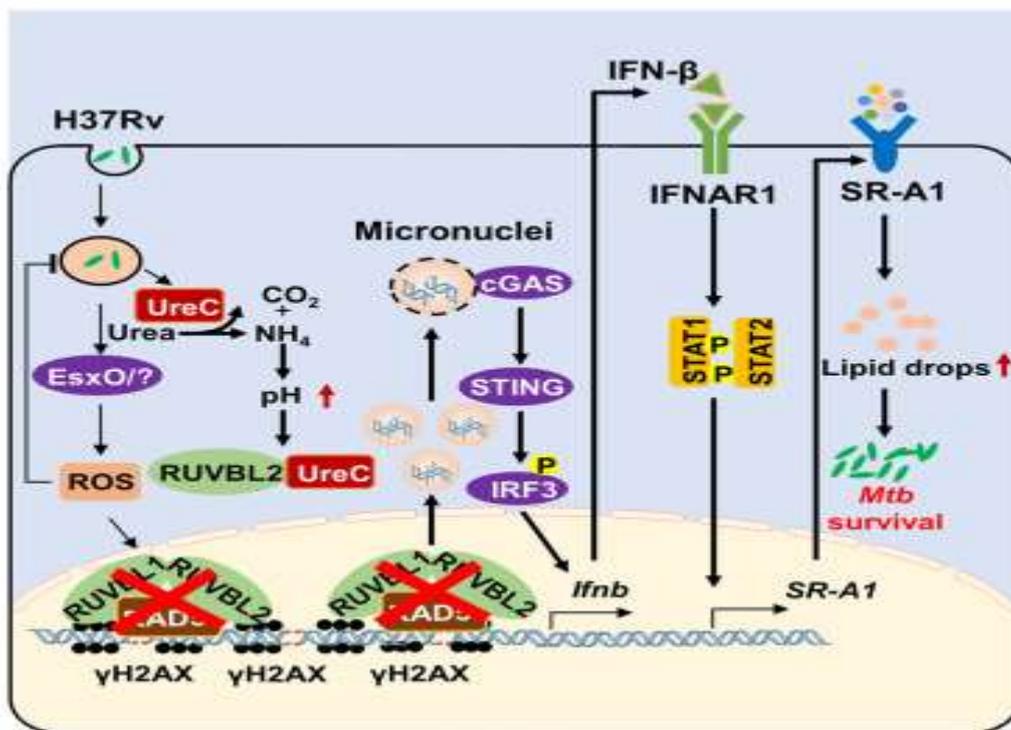
Madduri BTSa, Bell SL.

Bug in the code: TB blocks DNA repair.

Cell Host Microbe. 2023 Nov 8;31(11):1769-1771.

Proteggere il genoma della cellula è fondamentale per la sopravvivenza, ma l'infezione provoca danni che compromettono l'integrità genetica. In questo numero di Cell Host & Microbe, Lui et al. analizzano come il Mycobacterium tuberculosis sfrutta il danno al DNA utilizzando una proteina secreta che inibisce la riparazione del DNA per creare un ambiente favorevole alla replicazione batterica.

i micronuclei sono un indicatore di instabilità del genoma, ma UreC non ha indotto danni al DNA nei macrofagi di topo o nelle cellule umane in coltura. Invece, UreC ha inibito la riparazione del DNA interagendo con l'ATPasi RUVBL2 e impedendole di formare complessi funzionali di riparazione del DNA con i suoi partner RUVBL1 e RAD51.



Questa soppressione della riparazione del DNA ha portato alla formazione di micronuclei che hanno stimolato la via cGAS-STING citosolica di rilevamento del DNA, culminando nella produzione di interferone-β (IFN-β) che, attraverso un meccanismo sconosciuto, ha causato l'accumulo di goccioline lipidiche nei macrofagi .

Esperimenti di infezione nei topi con knockout condizionale e mieloide-specifico di RUVBL2 hanno ulteriormente supportato il meccanismo che è stato definito in vitro. Il legame di UreC a RUVBL2 dipendeva dall'attività catalitica di UreC; il trattamento di topi infetti da Mtb con acido acetidrossamico, un inibitore dell'ureasi utilizzato per trattare le infezioni del tratto urinario e i calcoli renali, ha ridotto la carica batterica nei polmoni.

Pertanto, oltre a indurre danni al DNA, Mtb inibisce anche la riparazione del DNA, portando a una segnalazione immunitaria innata indotta da danni al DNA che promuove la formazione di cellule schiumose.