

Il modo giusto di preoccuparsi

di Daron Acemoglu

Pubblicato il 14 maggio 2021 su Project Syndicate
Traduzione di Matteo Negrini

Il regno dei dinosauri si estinse 65 milioni di anni fa per un asteroide che si schiantò nell'attuale città di Chicxulub, in Messico.

Sebbene questo pezzo di roccia e metallo non fosse particolarmente grande - probabilmente circa dieci chilometri di diametro - colpì la Terra a più di 60.000 km/h, causando un'esplosione che fu miliardi di volte maggiore di quella della bomba atomica di Hiroshima e che distrusse ogni forma di vita nel raggio di 1.000 chilometri.

Ancora più grave fu il fatto che l'esplosione mandò negli strati superiori dell'atmosfera un'enorme nuvola di polvere e cenere che bloccò i raggi del sole per anni.

Questo impedì la fotosintesi e ridusse drasticamente le temperature, ragion per cui gli scienziati ritengono che fu quella polvere atmosferica e gli aerosol di solfato che, alla fine, uccisero i dinosauri e molte altre specie.

Se oggi un asteroide o una cometa simile si schiantasse sulla Terra, causerebbe un altro evento di estinzione di massa, spazzando via la maggior parte delle specie e la civiltà umana così come la conosciamo.

Questa - remota - possibilità è un esempio di quello che viene detto "*rischio esistenziale naturale*": un evento non causato dall'uomo che porta all'estinzione - o quasi - della nostra specie.

Vi sono, però, anche rischi esistenziali di origine antropica, creati dall'uomo.

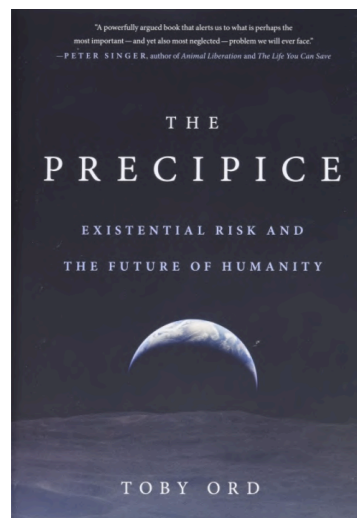
Come sostiene il filosofo dell'Università di Oxford Toby Ord nel suo nuovo stimolante libro, *The Precipice: Existential Risk and the Future of Humanity* (Bloomsbury, London, 2020), sono questi i rischi di cui dovremmo preoccuparci maggiormente, adesso e nel prossimo secolo.

Rischio e ricompensa

Ord riconosce che la scienza e la tecnologia sono gli strumenti più potenti che ha l'umanità per risolvere i problemi e raggiungere la prosperità. Ma ci ricorda che esistono sempre dei rischi associati ad un tale potere, specialmente se messo nelle mani sbagliate oppure se esercitato senza pensare alle conseguenze nel lungo termine.

Nello specifico, Ord sostiene che il rischio esistenziale antropogenico abbia raggiunto un livello allarmante, visto che, pur avendo sviluppato strumenti capaci di distruggere l'umanità, non abbiamo ancora acquisito la saggezza necessaria per riconoscere il pericolo in cui ci troviamo.

Egli ricorda che l'eminente astronomo del ventesimo secolo Carl Sagan lanciò un analogo allarme nel suo libro "*Pale Blue Dot*" del 1994, quando scrisse:





«In effetti, molti dei pericoli che affrontiamo derivano dalla scienza e dalla tecnologia; in particolar modo perché siamo diventati potenti senza diventare altrettanto saggi. La capacità di modificare il mondo che la tecnologia ci ha consegnato richiede oggi un grado di attenzione e lungimiranza come non ci è mai stato chiesto prima.»

Per Ord, questo divario tra potere e saggezza potrebbe decidere il futuro dell'umanità. Se, da una parte, potremmo scomparire dalla faccia della terra o subire un disastro tale da cancellare la maggior parte dei segni distintivi della nostra civiltà (dai vaccini e gli antibiotici all'arte e la scrittura), dall'altra, Ord vede nell'umanità il potenziale per una prosperità a lungo termine su scala cosmica: con saggezza e ingegnosità tecnologica, gli esseri umani potrebbero sopravvivere a questo pianeta e lanciare nuove civiltà attraverso lo spazio.

Questa visione di prosperità nel lungo periodo incide pesantemente sul giudizio di Ord, che ammette che potrebbero non esistere altre forme di vita intelligenti nell'universo. Se fossimo davvero soli, un'estinzione di massa che spazzasse via tutti quanti su questo pianeta eliminerebbe ovunque la possibilità di un'esistenza intelligente e consapevole.

Partendo da questo ragionamento, Ord arriva a quello che matematici ed economisti definirebbero un "ordinamento lessicografico delle preferenze".

In una situazione in cui vi sono più criteri da valutare, un ordine lessicografico assegna un'importanza preponderante ad uno di questi, in modo da semplificare il confronto tra due opzioni.

Ad esempio, in un ordine lessicografico tra cibo e riparo, sarebbe sempre da preferire qualunque opzione offra più cibo, indipendentemente da quanto più riparo offra l'altra possibilità.

La posizione filosofica di Ord equivale a un ordine lessicografico perché implica che dovremmo ridurre il rischio esistenziale al minimo, a qualunque costo.

Per Ord, un futuro in cui è stato minimizzato il rischio esistenziale deve prevalere su qualsiasi futuro in cui non lo sia stato, indipendentemente da qualunque altra considerazione.

Dopo aver stabilito questa gerarchia di base, Ord procede quindi con una panoramica dei diversi tipi di rischio esistenziale antropogenico, arrivando alla conclusione che la minaccia più grande possa provenire da una "superintelligenza" artificiale evolutasi al di fuori del nostro controllo.

Quando il progresso non è progresso

Si può far risalire il primo rischio esistenziale dipendente dalla scienza al momento in cui diventammo capaci di controllare le reazioni nucleari a catena e, di conseguenza, potemmo incominciare a realizzare armi atomiche.

Probabilmente, Ord ha ragione quando afferma che la nostra saggezza (sociale) non sia aumentata da quel fatidico momento, che vide il suo compimento nei bombardamenti di Hiroshima e Nagasaki.

Sebbene siano state create istituzioni, strumenti di regolamentazione, norme e altri meccanismi di internazionalizzazione che ci garantiscano da usi impropri della scienza, non si può certo dire che siano sufficienti.

Ord suggerisce che l'inadeguato quadro istituzionale odierno possa essere un fenomeno temporaneo, da affrontare a tempo debito, sempre che si riesca a sopravvivere al prossimo secolo o giù di lì.

«Siamo in un momento cruciale nella storia della nostra specie» scrive.

«Alimentato dal progresso tecnologico, il nostro potere è diventato così grande che, per la prima volta nella lunga storia dell'umanità, siamo in grado di distruggere noi stessi.»

È così che, nello scrivere il suo libro, Ord *«vorrebbe iniziare a colmare il divario tra la nostra saggezza e il nostro potere, permettendo all'umanità di avere una chiara visione di ciò che è in gioco, in modo che si possano fare le scelte necessarie per salvaguardare il nostro futuro.»*



Purtroppo, temo che, nella realtà, ciò non sia possibile, così come non scorgo nella nostra società e nei nostri leader alcun segno di saggezza capace di frenare il potere distruttivo della tecnologia.

A dire il vero, si potrebbe sostenere l'ottimismo di Ord con quello che il sociologo tedesco Norbert Elias definiva "*processo di civilizzazione*".

Secondo Elias, il processo di sviluppo economico e l'emergere, sin dal Medioevo, di istituzioni statali preposte a risolvere i conflitti e a controllare la violenza hanno portato all'adozione di modi e comportamenti che hanno reso possibile la convivenza nelle società di massa.

L'ipotesi di Elias sul perché le persone nelle economie avanzate siano diventate meno violente e più tolleranti è stata, recentemente, resa popolare dallo psicologo cognitivo e linguista dell'Università di Harvard Steven Pinker, nel suo libro "*The Better Angels of Our Nature: The Decline of Violence in History and Its Causes*".

Entrambi gli autori spiegano perché dovremmo continuare a volere un rafforzamento delle norme e delle istituzioni atte a controllare gli usi impropri di scienza e tecnologia.

Un tale *processo di civilizzazione* dovrebbe agire sui comportamenti individuali e, più in generale, sui rapporti sociali, ma non sembra che abbia influenzato molti leader politici, scienziati e tecnologi.

Per dirne una: il *processo di civilizzazione* avrebbe dovuto essere al suo apice nella prima metà del ventesimo secolo; eppure il chimico vincitore del Premio Nobel Fritz Haber fu entusiasta di usare le sue conoscenze scientifiche per creare e poi vendere armi chimiche all'esercito tedesco, durante la Prima Guerra Mondiale.

L'impatto del *processo di civilizzazione* non è stato particolarmente evidente neppure nel pensiero dei leader americani che ordinarono gli attacchi a Hiroshima e Nagasaki, così come negli atteggiamenti di altri leader politici che, dopo la Seconda Guerra Mondiale, abbracciarono con entusiasmo l'idea delle armi nucleari.

Alcuni potrebbero trovare motivo di speranza nel fatto che negli ultimi 75 anni non vi siano state altre guerre mondiali, ma questa visione ottimista non tiene conto delle molte volte in cui si è sfiorata la tragedia, non ultima la crisi dei missili cubani nel 1962 (l'episodio con cui Ord apre il suo libro).

Potrei portare molti altri esempi che contraddicono l'idea che stiamo diventando più "civilizzati", figuriamoci riuscire a controllare i rischi antropogenici o coltivare una saggezza collettiva!

Tutt'al più, continueremo ad affannarci per tenere d'occhio i nostri cattivi comportamenti e per adattarci ai continui cambiamenti dovuti alle scoperte scientifiche e all'innovazione tecnologica.

Questo solleva dubbi sulla validità del ragionamento di Ord.

Perché mai il tentativo di eliminare i futuri rischi esistenziali dovrebbe avere la priorità sugli sforzi tesi ad alleviare i mali e le sofferenze che le nostre scelte attuali stanno generando adesso e nel breve termine?

Supponiamo, paradossalmente, di poter abbassare significativamente le probabilità di estinguerci riducendo in schiavitù la maggior parte dell'umanità per i prossimi secoli.

Secondo l'ordinamento lessicografico di Ord, dovremmo scegliere questa opzione perché riduce al minimo il rischio esistenziale, pur preservando la possibilità per l'umanità di prosperare pienamente in un lontano futuro.

Non tutti sarebbero convinti da questo tipo di argomentazione. Io sono fra questi.

L'era delle macchine demoniache?

Per chiarire ulteriormente, si consideri il principale rischio esistenziale su cui si concentra Ord: l'uso potenzialmente scorretto dell'Intelligenza Artificiale (IA).

Ord stima che vi sia una possibilità su dieci che nei prossimi 100 anni l'umanità sia vittima di una malvagia "*superintelligenza*" (che chiama, eufemisticamente, "*IA non allineata*").



Per contro, reputa che il rischio esistenziale per l'umanità dovuto al cambiamento climatico sia pari ad uno su 1.000 e ad uno su 1.000.000 quello di una collisione con asteroidi o comete.

Sebbene molti altri esperti non assegnerebbero così alte probabilità alla minaccia della “*superintelligenza*”, Ord non è il solo a preoccuparsi delle ripercussioni a lungo termine della ricerca sull'IA. In effetti, tali preoccupazioni accomunano molti luminari della tecnologia, da Stuart Russell, dell'Università di Berkeley, al fondatore di Microsoft, Bill Gates, e a quello di Tesla, Elon Musk.



Matrix (1999) diretto da The Wachowskis, produzione Warner Bros.

In ogni caso, tutte queste personalità ritengono che - rischi esistenziali a parte - l'intelligenza artificiale porterà molti benefici. Pur essendo abbastanza informato su questi dibattiti per sapere che anche quest'ultima prospettiva è, in realtà, piuttosto traballante, la presa di posizione lessicografica di Ord lo porta a non considerare la maggior parte dei rischi *non* esistenziali correlati all'IA.

Restringendo l'ambito delle nostre analisi ad un contesto più vicino a noi, questa valutazione delle priorità mostra alcune problematiche. A mio parere, infatti, le probabilità che in tempi brevi possa emergere una “*superintelligenza*” sono basse e più basso ancora è il rischio che una “*superintelligenza malvagia*” possa distruggere la nostra civiltà.

Preferirei, quindi, che il dibattito pubblico si concentrasse maggiormente sui problemi che l'Intelligenza Artificiale sta creando già ora all'umanità, piuttosto che su possibili rischi futuri, intriganti ma improbabili.

Tornando ad oggi

Come ho già sostenuto qui e altrove, la traiettoria attualmente seguita dalla progettazione e dall'implementazione dell'IA ci sta portando fuori strada, causando una vasta gamma di problemi immediati.

Ben lungi dall'essere inevitabili o dal riflettere una qualche logica intrinseca della tecnologia, questi problemi evidenziano le scelte fatte (e imposteci) dalle grandi aziende tecnologiche e, in particolare, da un piccolo gruppo di dirigenti, scienziati e tecnologi all'interno di queste aziende (o nella loro orbita).

Uno dei guai più evidenti tra quelli causati dall'Intelligenza Artificiale è l'incessante processo di automatizzazione che sta progressivamente rimpiazzando i lavoratori, aumentando le disuguaglianze e sollevando lo spettro di una futura disoccupazione per ampie fasce della forza lavoro.

Non solo, questa ossessione per l'automazione è andata anche a scapito di una crescita della produttività, dato che ha portato dirigenti e scienziati a trascurare utilizzi dell'innovazione tecnologica che fossero più proficui e, allo stesso tempo, compatibili col lavoro umano.



Oggi, l'Intelligenza Artificiale viene usata in molti modi discutibili, nessuno dei quali capace di ispirare grandi speranze sul progresso morale dell'umanità.

Il funzionamento del sistema democratico, ad esempio, è stato contaminato non solo dall'esplosione di una disinformazione algoritmicamente amplificata, ma anche da nuove tecnologie di Intelligenza Artificiale che consentono a governi ed aziende di sorvegliare e manipolare i comportamenti di miliardi di persone.

Questa situazione rappresenta un duplice smacco.

Il controllo democratico, infatti, è il principale strumento a disposizione della società per tenere a freno eventuali comportamenti scorretti da parte delle *élite* politiche ed economiche, ma è proprio questo strumento ad esser stato danneggiato.

Una volta compromesso il processo democratico, non abbiamo modo di chieder conto alle *élite* dei danni che stanno causando: come faremo, allora, a risolvere le attuali difficoltà?

Possiamo, comunque, fare qualcosa.

I costi che l'IA ci sta infliggendo possono essere affrontati perché, a differenza dei rischi esistenziali su cui si concentra Ord, sono tangibili e facilmente riconoscibili. Prima, però, dobbiamo richiamare maggiore attenzione sul problema, in modo da fare pressione sui governi e sulle aziende affinché riconoscano i rischi che si stanno materializzando proprio ora.

Del resto, un settore tecnologico ripiegato sull'automazione e sulla manipolazione e sorveglianza antidemocratiche non è certo una buona base da cui partire nemmeno per affrontare i rischi di lungo periodo.

Nonostante non si debbano scartare a priori ipotetici rischi futuri per l'umanità, non possiamo permetterci di ignorare le minacce che sono proprio davanti a noi.

Può darsi che, ad un certo punto, arriveremo davanti a un precipizio, ma già adesso stiamo scivolando giù per una brutta china.

Daron Acemoglu insegna economia al MIT (*Massachusetts Institute of Technology*).

Vincitore nel 2005 della John Bates Clark Medal, è attualmente fra i cinque economisti più citati nel mondo.