

Il nuovo fuoco: un estratto dal libro di Ben Buchanan e Andrew Imbrie

di Ben Buchanan, Andrew Imbrie

24-10-2024

In *Il nuovo fuoco. Guerra, pace e democrazia nell'era dell'intelligenza artificiale* - edito da Bocconi University Press con prefazione di Alessandro Aresu - Ben Buchanan e Andrew Imbrie riflettono sulla geopolitica del futuro e sul ruolo della democrazia alla luce dei possibili impatti generati dallo sviluppo delle tecnologie basate sull'intelligenza artificiale.

Buchanan e Andrew Imbrie sono stati tra i pensatori che hanno contribuito all'Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence firmato dal presidente Joe Biden, il loro volume rappresenta quindi anche un'occasione per confrontarsi con l'azione più recente degli Stati Uniti in quest'ambito. Pubblichiamo di seguito, per gentile concessione dell'editore Bocconi University Press / Egea, un estratto del testo. Copyright © 2022 Benjamin Louis Buchanan and William Andrew Imbrie.

Il fuoco comincia da una piccola scintilla. Questa può nascere da due bastoncini sfregati tra loro, da un fiammifero acceso gettato incautamente per terra, da un guasto elettrico o da un'infinità di inneschi innocui ma imprevedibili. La scintilla si accende e, con il giusto carburante, cresce in maniera esponenziale. Ciò che è cominciato come una fiamma quasi impercettibile può trasformarsi velocemente in una vampata e poi deflagrare in un incendio incontrollato che devasta il paesaggio e distrugge tutto ciò che incontra sulla sua strada. Il potere del fuoco non è frutto solo del suo calore e del suo fumo, ma anche della sua accelerazione in velocità e forza; se lasciato incontrollato, si scatena e infuria fino a quando non troverà più nulla da bruciare.

Eppure, nonostante la forza distruttiva di cui è capace, il fuoco è anche la base della civiltà. Sin dagli albori, la capacità dell'umanità di accendere, controllare e fermare il fuoco è stata fondamentale per riscaldare le caverne e per tutto ciò che è venuto dopo. Gran parte delle invenzioni più importanti dell'umanità - la forgiatura dei metalli, la propulsione a vapore, la lavorazione del vetro e l'elettricità - deriva da questa abilità. Secondo alcuni scienziati, lo stesso sviluppo del cervello umano sarebbe dipeso dalla cottura della carne sul fuoco, che ha consentito al cibo di essere digerito con un minor dispendio di energia[1]. Gli esseri umani hanno imparato a sfruttare la potenza esponenziale del fuoco in moltissimi modi, e per farlo hanno dovuto capire come utilizzare la sua forza a fin di bene e come domarne le insidie.

Ma l'umanità ha utilizzato anche la forza distruttiva del fuoco. L'Impero bizantino la brandì con grande successo militare, prima durante l'assedio di Costantinopoli nel 672 d.C. e poi nei secoli successivi. In battaglia, le truppe bizantine sparavano contro i nemici un composto appositamente formulato, in grado di bruciare anche a contatto con l'acqua[2]. Quando il composto colpiva il bersaglio, la potenza del fuoco si attivava incendiando l'equipaggiamento nemico e mettendo in fuga i soldati. Da quel momento in poi le fiamme della guerra sono diventate sempre più letali.

Esiste una forza altrettanto proficua e pericolosa? Una forza altrettanto legata a doppio filo alla crescita esponenziale dei suoi componenti principali? Benvenuti nell'era dell'intelligenza artificiale. Scegliendo la metafora del fuoco per spiegare questa nuova era, respingiamo l'affermazione più diffusa e ottimistica che vede l'intelligenza artificiale come «la nuova elettricità»[3]. Nella vita

moderna l'elettricità è ovunque: è talmente onnipresente e pervasiva nella nostra società che diamo per scontata la sua sicurezza. Senza il benché minimo timore, viviamo e lavoriamo fianco a fianco con gli enormi cavi brulicanti di energia che attraversano i nostri quartieri e alimentano le nostre vite. Grazie alla professionalizzazione delle attività di generazione, trasmissione e utilizzo dell'elettricità, quasi tutte le moderne interazioni umane con la corrente elettrica sono benefiche. La metafora dell'elettricità vorrebbe l'intelligenza artificiale destinata a seguire lo stesso percorso, trasformando la società in maniera altrettanto profonda e altrettanto positiva.

Non è escluso che un giorno si arrivi a un epilogo di questo tipo, ma al momento questa visione pecca di ottimismo. La nostra domestichezza con l'intelligenza artificiale non è neanche lontanamente paragonabile a quella che abbiamo raggiunto con l'elettricità[4]. Oggi il nostro incontro con l'intelligenza artificiale assomiglia a quello che i nostri lontani antenati ebbero con il fuoco. Se sapremo gestire bene questa tecnologia, essa diventerà una forza possente per il bene globale, aprendo la strada a molte invenzioni trasformative. Se la adopereremo troppo velocemente e senza la dovuta lungimiranza, l'intelligenza artificiale avvamperà in modi che non riusciremo a controllare. Se la utilizzeremo per distruggere, metterà nelle mani dei governi più forti della Terra armi più potenti da brandire in una competizione geopolitica ad alto grado di infiammabilità. Il fatto che la frequente analogia con l'elettricità neghi questa ampia gamma di possibili esiti non ci rende più sicuri, ma solo meno preparati.

Sono tre le scintille che accendono il nuovo fuoco: i dati, gli algoritmi e la potenza di calcolo. Gli attuali sistemi di intelligenza artificiale utilizzano la potenza di calcolo per eseguire algoritmi che istruiscono le macchine su come imparare dai dati. La crescita esponenziale delle dimensioni degli insiemi di dati, delle capacità degli algoritmi e della potenza dei computer ha consentito all'intelligenza artificiale di compiere progressi che hanno lasciato stupefatti anche gli osservatori più scettici.

L'intelligenza artificiale ha dimostrato capacità straordinarie in aree che vanno dal riconoscimento vocale e facciale alla generazione di video, la traduzione linguistica, la scrittura e altro ancora. Sono stati questi progressi a far entrare a pieno titolo l'intelligenza artificiale nelle nostre case e nelle nostre aziende. L'IA alimenta Siri e Alexa, ci dà consigli durante la navigazione sul web, ci aiuta a guidare per strade e autostrade e, silenziosamente, fa funzionare meglio di prima molti dei sistemi tecnologici che utilizziamo quotidianamente. Ancora più impressionante è l'attuale capacità dei sistemi di intelligenza artificiale di imitare qualità inerentemente umane, come l'immaginazione e l'intuizione. In rapida successione, questi sistemi hanno raggiunto traguardi - tra cui la vittoria in giochi che richiedono una pianificazione strategica tremendamente complessa e la soluzione di alcuni dei più difficili problemi scientifici - che gli esperti ritenevano distanti più di un decennio. È estremamente probabile che i prossimi anni abbiano in serbo progressi ancora più sorprendenti e potenti per l'intelligenza artificiale. Le nuove capacità dell'intelligenza artificiale sono allo stesso tempo meraviglie e distrazioni. Troppo spesso la discussione sull'IA si concentra su ciò che la tecnologia è in grado di fare, omettendo di soffermarsi sulle persone che la inventano, la perfezionano e la impiegano. Gli osservatori fanno previsioni conturbanti sull'avvento di una superintelligenza in grado di superare le capacità cognitive dell'umanità, ma si soffermano molto meno sulle modalità con cui l'intelligenza artificiale interagisce con la geopolitica[5]. Eppure, la storia ci insegna che, dalle carrozze ai cannoni, dagli aeroplani alle bombe atomiche, più importanti dell'innovazione in sé sono il come e il perché le persone la adoperano[6]. Qualsiasi indagine sull'intelligenza artificiale dovrebbe porre l'enfasi su di noi: su come adoperiamo il nuovo fuoco e a quale scopo. Questo libro parla di scelte umane. Le decisioni che governi, aziende e università prendono in merito all'intelligenza artificiale saranno quelle che più incideranno su questo secolo appena iniziato, sia dal punto di vista tecnologico che geopolitico. L'intelligenza artificiale ci impone di rispondere di nuovo ad alcune domande fondamentali sull'umanità: come governiamo? Come

combattiamo? E l'interrogativo probabilmente più importante: per cosa combattiamo? Per capire cosa fare con questa nuova tecnologia, illustriamo tre punti di vista confliggenti e talvolta sovrapposti: quelli di evangelisti, guerrieri e Cassandre.

Gli evangelisti ritengono che l'umanità sia chiamata a prendersi cura del nuovo fuoco per il bene comune. Secondo i sostenitori di questo punto di vista l'intelligenza artificiale è in grado di trasformare la civiltà umana in meglio e l'uso dell'IA come arma bellica e geopolitica rischia di distogliere l'attenzione da tutto ciò che di buono può fare per l'umanità. Nelle file degli evangelisti ci sono molti ricercatori e pionieri dell'intelligenza artificiale mossi dal desiderio di «risolvere l'intelligenza» per svelare finalmente alcuni dei misteri più spinosi della scienza. In un momento storico in cui la scoperta e l'invenzione scientifica sembrano più complicate che mai, gli evangelisti vogliono costruire macchine in grado di vedere, creare, pianificare e aiutare l'umanità a raggiungere il suo pieno potenziale. Proprio come la scoperta del fuoco, le scoperte nell'IA - o ad opera dell'IA - regaleranno all'umanità nuovi standard di vita e nuovi gradi di comprensione. Se la gestiremo bene, sostengono gli evangelisti, la forza esponenziale del nuovo fuoco potrà creare progresso per tutti e offrire a tutti vantaggi a somma positiva.

I guerrieri, dal canto loro, vogliono sfruttare il nuovo fuoco non solo per scopi scientifici ma anche per la sicurezza nazionale. Se dai tempi dei Bizantini quasi tutte le forze armate hanno cercato e sono riuscite a incendiare il nemico con una fiamma di qualche tipo, molti strateghi odierni pensano che l'intelligenza artificiale possa essere utile per vincere le guerre e come strumento di deterrenza. Essi ritengono che la competizione geopolitica non risparmi nessun ambito della vita umana e che le nazioni si comporteranno con l'intelligenza artificiale come con qualsiasi altra questione globale, dal cambiamento climatico alle pandemie: agiranno per il proprio tornaconto, perlopiù in competizione e solo raramente in cooperazione tra loro. Secondo i guerrieri, sono soprattutto le democrazie a dover imparare a maneggiare il nuovo fuoco per non rischiare di assistere dalle retrovie all'uso che le autocrazie faranno dell'intelligenza artificiale per accentrare il controllo in patria ed estendere la propria influenza a livello globale. Essere lenti ad adattarsi significherà cedere il vantaggio ad altri. Contemplare la necessità di impiegare l'intelligenza artificiale per la sicurezza e per ricavarne vantaggi geopolitici può essere inquietante, ammettono i guerrieri, ma ignorarla sarebbe da ingenui. A volte il mondo è più a somma zero di quanto siamo disposti ad ammettere. Le Cassandre sono spaventate dal nuovo fuoco. Cassandra era la sacerdotessa troiana le cui profezie erano corrette ma venivano puntualmente ignorate. Le odierne Cassandre sostengono che l'intelligenza artificiale è meno utile e più pericolosa di quanto previsto da molti dei suoi fautori. Se i guerrieri temono che le democrazie siano troppo lente nell'applicare l'intelligenza artificiale alla sicurezza nazionale, le Cassandre paventano invece che le nazioni possano farlo troppo velocemente. Temono le debolezze intrinseche dell'IA così come la sua capacità di essere sia potente che inabile, e prevedono che i sistemi di intelligenza artificiale falliranno catastroficamente nei momenti di cruciale importanza. Se impiegata troppo in fretta e senza adeguate accortezze e vincoli etici, l'IA potrebbe avere effetti enormemente distruttivi e a somma negativa, non alla stregua di un Terminator vendicativo ma come una vampata che investe tutto con il suo furore privo di emozioni.

Nel libro ci serviamo dei punti di vista di evangelisti, guerrieri e Cassandre per analizzarne i presupposti divergenti e le diverse enfasi, ma a volte i confini tra queste categorie si confondono. Alcuni osservatori rimangono della stessa idea su molte delle questioni che intersecano intelligenza artificiale e geopolitica, mentre altri cambiano punto di vista a seconda della tematica. Un evangelista può preoccuparsi dei rischi dell'intelligenza artificiale così come un guerriero o una Cassandra possono credere nei potenziali benefici della tecnologia per il bene comune.

Tutte e tre le prospettive hanno i loro meriti. L'intelligenza artificiale è tecnicamente impressionante ed è probabile che riesca a contribuire a scoperte che renderanno la vita migliore e più appagante per gran parte dell'umanità. Allo stesso tempo, l'intelligenza artificiale è già diventata una linea di

faglia tra Stati Uniti e Cina e, più in generale, tra le forze della democrazia e quelle dell'autocrazia. Entrambe le superpotenze investono ogni anno miliardi di dollari nell'IA non solo per far progredire l'umanità ma anche per ricavarne un vantaggio economico e militare sull'altra. Questa competizione, proprio come il nuovo fuoco, non fa che accelerare. I guerrieri hanno ragione a volere che le democrazie vincano, ma le Cassandre hanno ragione a mettere in guardia da una corsa al ribasso con effetti deleteri sui diritti umani, sull'affidabilità della tecnologia e sulla sicurezza internazionale. L'interazione tra questi tre punti di vista determinerà ciò che l'intelligenza artificiale può fare, chi ne trarrà beneficio e quali saranno le ricadute su tutti noi. Nelle pagine che seguono daremo voce agli esponenti di ciascuno dei tre punti di vista. Alcuni di essi spingono sempre più in là la frontiera tecnologica dell'intelligenza artificiale facendole compiere progressi un tempo impensabili. Altri sfruttano queste invenzioni per promuovere gli interessi delle loro nazioni, siano esse democratiche o autocratiche. Altri ancora invitano alla cautela, avvisandoci che stiamo oltrepassando i confini etici o creando conseguenze indesiderate. A volte le Cassandre riescono a far sentire la propria voce e a ottenere un'attenuazione dei rischi, ma la spinta all'accelerazione della competizione geopolitica e della ricerca sull'intelligenza artificiale è inarrestabile. L'intelligenza artificiale plasmerà l'arte di governo, e l'arte di governo plasmerà l'intelligenza artificiale.

[1] Karina Fonseca-Azevedo, Suzana Herculano-Houzel, «Metabolic Constraint Imposes Tradeoff between Body Size and Number of Brain Neurons in Human Evolution», in *Proceedings of the National Academy of Sciences of the United States of America*, 109, n. 45, 6 novembre 2012, pp. 18571-18676.

[2] John Pryor, Elizabeth M. Jeffreys, «The Age of the DROMWN: The Byzantine Navy 500-1204», in *The Age of the DROMWN*, Boston, Brill, 2006, pp. 607-609.

[3] Catherine Jewell, «Artificial Intelligence: The New Electricity», *WIPO Magazine*, giugno 2019; Andrew Ng, «Artificial Intelligence Is the New Electricity», *Stanford Graduate School of Business MSx Future Forum*, 25 gennaio 2017, YouTube video, 1:27:43.

[4] Ernest Freeberg, *The Age of Edison: Electric Light and the Invention of Modern America*, London, Penguin, 2013, p. 9. Per inciso, c'è chi, come Sundar Pichai di Google, usa l'elettricità e il fuoco come metafore intercambiabili, ma senza fare riferimento all'ampia gamma di scenari associati al fuoco.

[5] Ray Kurzweil, *The Singularity Is Near: When Humans Transcend Biology*, New York, Viking Press, 2005; trad. it. *La singolarità è vicina*, Rimini, Maggioli Editore, 2008. A nostro avviso, l'intelligenza va intesa non come il superamento di una singola soglia ma piuttosto come la compresenza di uno spettro di possibilità. Da un lato, l'intelligenza è un parametro dipendente dal compito assegnato, sul quale il confronto con gli esseri umani non sempre serve a gettare luce: sappiamo che l'intelligenza artificiale eccelle in alcuni compiti, come il Go, e che fa molta più fatica in altri, come la sciarada. Dall'altro, l'intelligenza è caratterizzata dalla capacità di generalizzare con facilità e abilità a partire da più compiti e di apprendere un'ampia gamma di nuovi compiti con flessibilità. Shane Legg, Marcus Hutter, «A Collection of Definitions of Intelligence», arXiv, 15 giugno, 2007; François Chollet, «On the Measure of Intelligence», arXiv, 5 novembre 2019.

[6] David Edgerton, *The Shock of the Old: Technology and Global History since 1900*, Oxford, Oxford University Press, 2011; William H. McNeill, *The Pursuit of Power: Technology, Armed Force, and Society since A.D. 1000*, Chicago, University of Chicago Press, 2013; trad. it. *Caccia al potere*, Milano, Feltrinelli, 1984.