DOPPIOZERO

Stephen Hawking e l’intelligenza artificiale

Giuseppe O. Longo

18 Marzo 2018

La morte del grande astrofisico britannico Stephen Hawking ha avuto una vastissima eco anche sulla stampa non specializzata, che ha ricordato, in termini generici ma suggestivi, le conquiste scientifiche del grande studioso. Non c’è dubbio che alla sua eccezionale popolarità abbia contribuito anche la gravissima invalidità, che da un certo momento in poi gli ha impedito di muovere tutti i muscoli del corpo e perfino di parlare. Hawking riusciva a comunicare con l’esterno grazie a un raffinato software, dunque grazie all’intelligenza artificiale (IA). Tuttavia, nonostante il beneficio decisivo ottenuto da questa tecnologia, lo scienziato aveva una visione lucidissima non solo dei vantaggi ma anche dei rischi che potrebbe comportare uno sviluppo eccessivo dell’IA.

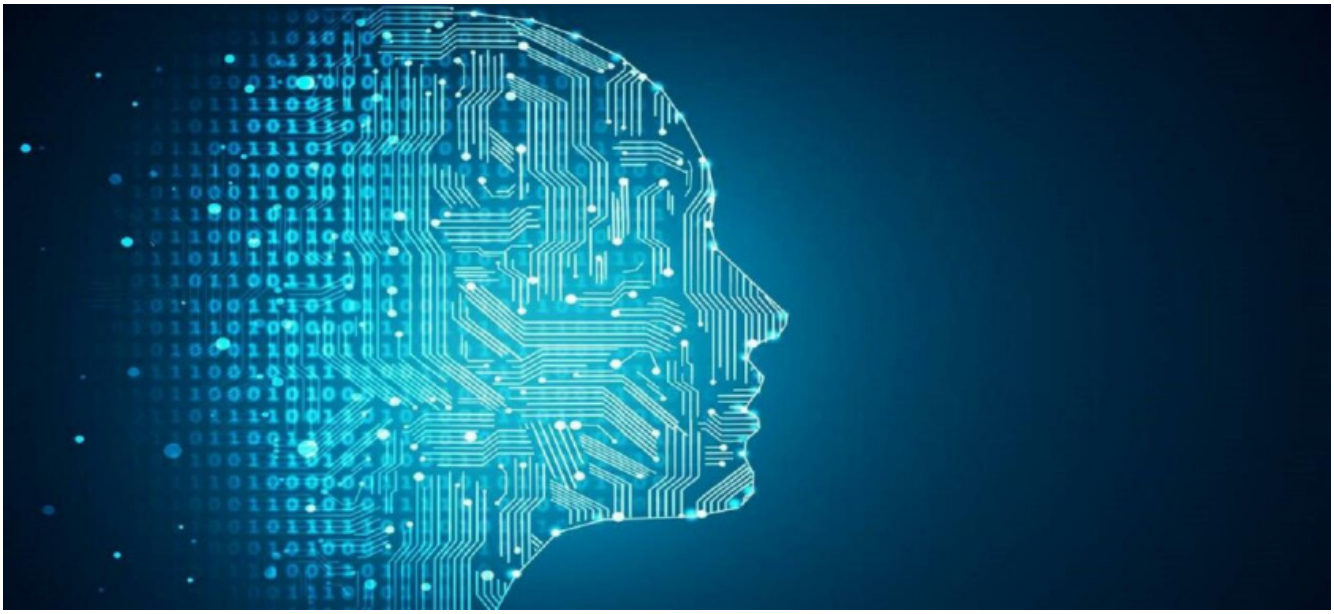
Già settant’anni fa il grande scienziato e filosofo Norbert Wiener, autore del pioneristico *Cybernetics or Control and Communication in the Animal and the Machine* (1948), aveva segnalato i pericoli derivanti dalla costruzione di macchine (troppo) intelligenti:

Se vorremo costruire macchine capaci di apprendere e di modificare il comportamento in base all’esperienza, dovremo accettare il fatto che ogni grado di indipendenza fornito ad esse potrebbe produrre un ugual grado di ribellione nei nostri confronti. Una volta uscito dalla bottiglia, il genio non avrà alcuna voglia di ritornarci, e non c’è motivo di aspettarsi che le macchine siano ben disposte verso di noi. In breve, solo un’umanità capace di rispetto e deferenza sarà capace di dominare le nuove potenzialità che ci si aprono davanti. Possiamo adottare un atteggiamento umile e condurre una vita buona con l’ausilio delle macchine, oppure possiamo adottare un atteggiamento arrogante e perire.

Questa posizione cauta, improntata a una grande sensibilità etica, va confrontata con la posizione di un altro grande studioso, il gesuita Pierre Teilhard de Chardin, che qualche anno dopo così si espresse:

Il fine della Natura è la sua fine. L'umanità si è per troppo tempo limitata all'empirismo docile e alla rassegnazione paziente. È giunta l'ora di dominare la Natura, di farla parlare, di domarla, d'inaugurare una fase nuova, nel corso della quale l'intelligenza si rivolgerà contro l'Universo dal quale è nata per correggerlo, rinnovarlo, fargli rendere sino in fondo tutto ciò che può fornire alla sua porzione cosciente come accrescimento di felicità e attività.

Quella di Chardin è una posizione progressista, quasi trionfalistica, che culmina nella visione del Punto Omega, uno stato asintotico in cui tutte le intelligenze umane e artificiali si fonderanno per dar luogo a un'intelligenza cosmica capace di perfezionare la creazione stessa, collaborando attivamente con il Creatore. All'opposto, la previsione che il fine della Natura sia la fine della Natura è vista con la massima preoccupazione dal filosofo tedesco Hans Jonas, autore del *Principio responsabilità* (1979), secondo il quale "La Natura non poteva correre rischio maggiore di quello di far nascere l'uomo. Nell'uomo la Natura ha distrutto se stessa."



Resta il fatto che la costruzione da parte nostra di artefatti sempre più progrediti, specie di tipo cognitivo, provoca quella che Günther Anders ha chiamato “vergogna prometeica”, cioè il senso umiliante di inferiorità che provano gli umani rispetto agli strumenti da loro stessi costruiti, tanto più rapidi, precisi ed efficienti. Paradossalmente, questo senso di inadeguatezza spinge l’uomo ad affidarsi sempre più alla macchina, operando una “delega tecnologica” analoga alla delega specialistica che si opera da sempre nei confronti dei medici, degli avvocati, dei barbieri. La delega tecnologica rende gli umani sempre più inetti e tributari dell’artificiale. Al limite, si configura una dipendenza che asintoticamente porta alla schiavitù.

Tutto ciò si può trasferire in particolare nel settore dell’IA, che è uno dei più dinamici sia per l’impegno dei ricercatori sia per i profitti, presenti e previsti, delle aziende. È difficile immaginare quello che potrà fare in futuro un’entità dotata di IA superiore, capace quindi (è uno scenario, ma non è privo di verosimiglianza) di evolversi senza bisogno dell’intervento umano. Come dichiara Nick Bostrom nel suo libro *Superintelligenza. Tendenze, pericoli, strategie* (Bollati Boringhieri, 2018), non si deve cedere alla tentazione di immaginare una superintelligenza in termini antropomorfi: i concetti di “intelligente” e “stupido” si basano sulla nostra esperienza, ma le differenze di capacità intellettiva all’interno dei gruppi umani sono insignificanti rispetto alla differenza tra un essere umano qualsiasi e una ipotetica superintelligenza. Dice Bostrom:

Invece di immaginare che un’IA superintelligente sia come un genio scientifico in confronto a una persona comune, sarebbe più appropriato pensare che sia come una persona comune rispetto a un insetto o un verme.

Di fronte a questa prospettiva si ripropone la domanda di Wiener: siamo sicuri di poter governare una superintelligenza, una volta che l’abbiamo fatta uscire dalla bottiglia? A chi risponde fiducioso “basterà spegnere l’interruttore” si può obiettare che forse quelle macchine non avranno nessun interruttore… Lo scopo, non sempre dichiarato ma evidente, delle ricerche sull’IA è quello di costruire delle entità cognitive dotate di un’intelligenza ad ampio spettro, come quella umana. Una volta raggiunto questo scopo, quanto tempo occorrerà a queste macchine per superare e surclassare le nostre dotazioni intellettuali? Per Bostrom occorrerà poco, anzi pochissimo: basterà un piccolo incremento e quelle IA partiranno a razzo verso traguardi inimmaginabili e soprattutto inarrivabili da

parte nostra. Il fatto è che l'evoluzione tecnica, innescata dall'uomo, è molto più rapida dell'evoluzione biologica e quindi l'uomo sarà presto "obsoleto", per usare un'altra espressione di Anders: anzi per molti versi lo è già.

Secondo alcuni futurologi l'uomo potrà sfuggire a questo destino di destituzione solo salendo a cavalcioni di queste IA, cioè, in termini espliciti, ibridandosi con le macchine, ed è ciò che in parte sta già avvenendo. Ma le conseguenze saranno dirompenti, come hanno previsto per esempio Marvin Minsky e Hans Moravec: l'umanità come la conosciamo e come possiamo immaginarla nel futuro prossimo scomparirà e saranno i robot a ereditare la terra e a padroneggiarla.

Sulla base di queste fosche previsioni, nel gennaio 2015 Bostrom ha firmato una lettera aperta, che è stata sottoscritta da molti altri scienziati, tra cui appunto Stephen Hawking, per mettere in guardia politici, ricercatori e sociologi sui potenziali pericoli di uno sviluppo eccessivo dell'IA. Non tutti sono d'accordo con Bostrom e gli altri firmatari e sostengono posizioni molto più ottimistiche. È vero infatti che l'IA è un territorio ricco di promesse straordinarie: grazie ai suoi sviluppi l'umanità potrebbe migliorare la propria esistenza, conseguire traguardi oggi neppure immaginabili, vivere più a lungo e sconfiggere patologie invalidanti. Ma, e questa è la preoccupazione che Hawking condivise con Bostrom, le nostre creature superintelligenti potrebbero impadronirsi delle redini della loro evoluzione e, come accade ai figli ribelli nei confronti dei genitori, non prestare più orecchio alle nostre esortazioni, ignorandoci come residui anacronistici di un'epoca ormai tramontata. Del resto, non è impensabile che, giunta a un certo grado di sviluppo, alla civiltà, per progredire ulteriormente, gli uomini non solo non bastino, ma non siano neppure più necessari...

Insomma, ancora una volta si rivela la natura duplice della tecnica, che essa eredita dal suo inventore Prometeo, raffinato artigiano e insieme abile truffatore: la tecnica ci libera sì da molti dei ceppi dovuti al nostro retaggio umano, ma ci impone anche schiavitù nuove e impreviste. Saremo in grado di trovare in noi la forza di rinunciare agli sviluppi rischiosi dell'IA e di adottare quelli benefici? E avremo la saggezza che ci consenta di distinguere gli uni dagli altri?

Se continuiamo a tenere vivo questo spazio è grazie a te. Anche un solo euro per noi significa molto.
Torna presto a leggerci e SOSTIENI DOPPIOZERO

