

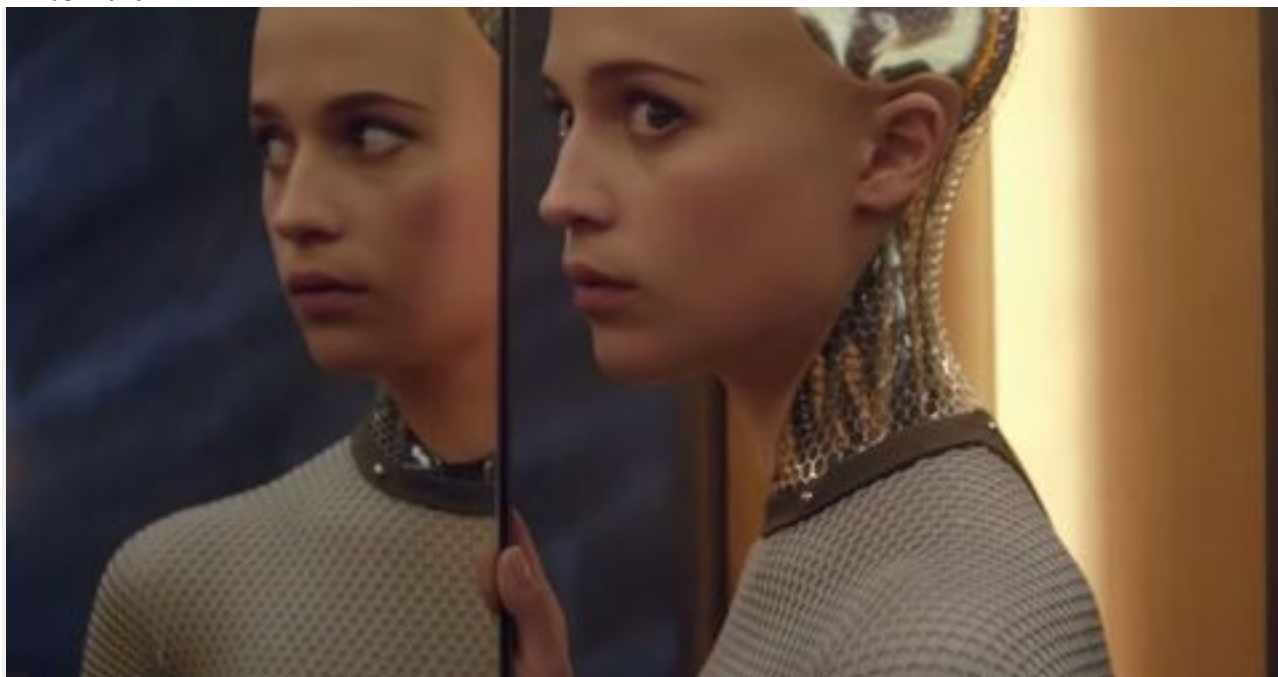
Esquire

L'intelligenza artificiale sarà la nostra ultima invenzione

Lo sostiene James Barrat nel suo libro: un viaggio nel mondo della ricerca sull'AI, che prima o poi sarà in grado di produrre sé stessa.

DI ROBERTO PAURA

11/03/2019



COURTESY UNIVERSAL

Eliezer Yudkovsky è uno di quei personaggi bizzarri in cui si può imbattere solo in quel laboratorio permanente della postumanità che è la Silicon Valley. Quando il giornalista e scrittore James Barrat lo va a incontrare, riceve per e-mail le coordinate per raggiungerlo col Gps, con la richiesta di non divulgarle a nessuno, e non riesce a farsi dare nemmeno un numero di telefono. Yudkovsky vive la sua vita pressoché esclusivamente su Internet. Rilascia raramente interviste, solo su Skype e per non più di mezz'ora. Comunica per mail, ma solo in modo stringato.

La sua fama è legata a un curioso esperimento, noto come AI-Box Experiment, svolto tra il 2002 e il 2005. È una versione più sofisticata (e più inquietante) del test di Turing, immaginato da Alan Turing per riuscire a riconoscere una vera intelligenza artificiale da un programma che imita semplicemente il comportamento umano senza capirlo. Nell'AI-Box

Experiment, Yudkovsky ricopriva il ruolo di un'ipotetica intelligenza artificiale generale (AGI, *artificial general intelligence*), vale a dire un'IA di intelligenza pari a quella umana (obiettivo non ancora raggiunto dallo sviluppo tecnologico). I suoi sfidanti, chiamati "Guardiani", avevano l'obiettivo di impedirgli di uscire dalla "scatola", ossia dal computer su cui gli ipotetici programmatori dell'IA lo avevano confinato per impedirgli di diffondersi nel mondo e provocare danni. Compito di Yudkovsky era di convincere i Guardiani a farlo "scappare" dal confinamento: se fosse riuscito a persuaderli, avrebbe vinto la scommessa.



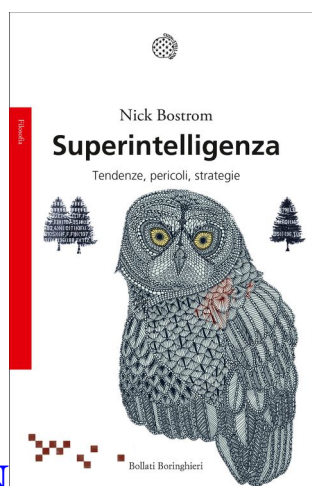
courtesy Nutrimenti

Leggenda vuole che a sfidarlo siano stati alcuni tra i principali magnati dalla Silicon Valley, tutti miseramente sconfitti e costretti a pagare puntate di milioni di dollari. La verità, secondo Yudkovsky, è più prosaica: fu sfidato cinque volte, vinse tre volte, e guadagnò solo poche centinaia di dollari. Ma l'esperimento funzionò: le sue tre vittorie convinsero i suoi sfidanti della facilità con cui un'IA di livello umano può liberarsi dalla sua originaria programmazione per perseguire i propri scopi (all'esperimento è ispirato il film *Ex Machina*). Se può essere così semplice per un uomo molto intelligente, quale difesa potremo mai immaginare contro una superintelligenza artificiale centinaia se non migliaia di volte superiore all'essere umano?

Il dibattito sui rischi dell'AGI e della superintelligenza artificiale è ormai giunto, con un certo ritardo, anche in Italia. Dopo l'uscita di *Superintelligenza* di Nick Bostrom e di *Vita 3.0* di Max

Tegmark (rispettivamente a capo del Future of Humanity Institute e del Future of Life Institute, i principali think-tank mondiali che lavorano sul tema dei rischi esistenziali dell'IA), arriva in italiano (edito da Nutrimenti nella traduzione di Daniela Pezzella e Monica Pezzella) anche il testo fondativo del genere, *La nostra invenzione finale* di James Barrat. A differenza dei precedenti, Barrat non è uno scienziato o un filosofo di professione, ma un giornalista e documentarista che – racconta il risvolto di copertina del libro – “inizialmente inebriato dal potenziale e dalle promesse dell'intelligenza artificiale, viene poi contagiato da un inquietante scetticismo riguardo al futuro roseo prospettato dai cacciatori di Asi (superintelligenza artificiale, ossia superiore a quella dell'uomo)”.

Il suo libro *Our Final Invention* ha avuto un'influenza enorme, trasformando questi temi di nicchia, all'epoca studiati solo da un eccentrico think-tank californiano, il Machine Intelligence Research Institute, fondato nel 2000 dall'ancor più eccentrico Eliezer Yudkovsky, in temi al centro del dibattito mondiale. Il libro di Barrat, e gli esperimenti di Yudkovsky, hanno messo abbastanza paura ai miliardari della Silicon Valley da spingerli a investire decine di milioni di dollari nelle strategie di sicurezza per contenere la superintelligenza (uno dei più generosi convertiti, com'è noto, è Elon Musk, che continua a staccare assegni da capogiro per i think-tank che si occupano di questi temi).



[COMPRA SU AMAZON](#)

courtesy Bollati Boringhieri

Non tutti, nella Silicon Valley, sono convinti che Barrat e gli IA-apocalittici abbiano ragione. Dall'altro lato della barricata, racconta il libro, ci sono quelli che l'autore definisce “singolaritaristi”, ossia gli avventisti della Singolarità tecnologica profetizzata dal tecnologo Ray Kurzweil per la metà di questo secolo: sarà il momento in cui il ritmo dell'accelerazione tecnologica raggiungerà la cuspide, portando alla nascita di una super-IA che cambierà completamente il nostro mondo, aprendo la strada alla

“società dell’abbondanza” (o della “post-scarsità”), un autentico paradiso in terra. A prescindere da chi abbia ragione, tuttavia, resta il fatto che l’AGI (o l’ASI) sarà davvero l’ultima invenzione dell’essere umano, come aveva profetizzato per primo il matematico Irving John Good.

I.J. Good è un altro dei singolari protagonisti del libro di Barrat. Morto nel 2009, lavorò alla decrittazione dei codici cifrati nazisti nel gruppo di Alan Turing e come lui si interessò per tutta la vita al problema dell’intelligenza artificiale. Nel 1965, nel suo articolo *Speculations Concerning the First Ultraintelligent Machine*, Good scrisse: “Poniamo che una macchina ultraintelligente si rivelasse di gran lunga più efficiente degli uomini più intelligenti in tutte le attività intellettive. Poiché la progettazione delle macchine rientra nelle attività intellettive, una macchina ultraintelligente potrebbe progettare macchine migliori; vi sarebbe, senza dubbio, una ‘esplosione di intelligenza’ e l’intelligenza dell’uomo ne uscirebbe sconfitta. Infatti la prima macchina ultraintelligente sarà l’ultima invenzione che l’uomo avrà bisogno di inventare, ammesso che la macchina sia abbastanza docile da dirci come tenerla sotto controllo”.

PUBBLICITÀ - CONTINUA A LEGGERE DI SEGUITO

La storia è tutta qui, risalente ad anni in cui personaggi come Bostrom, Musk o Yudkovsky non erano nemmeno nati. Good vide prima di tutti gli altri vantaggi e svantaggi della superintelligenza. Inizialmente ottimista, verso la fine della sua lunga vita cambiò idea. Nel 1998 scrisse, in una nota biografica redatta in terza persona: “*Speculations Concerning the First Ultraintelligent Machine* cominciava così: ‘La sopravvivenza dell’uomo dipende dalla precoce costruzione di una macchina ultraintelligente’. Queste le sue [di Good] parole durante la guerra fredda, ma egli sospetta che la parola ‘sopravvivenza’ sia da sostituire con ‘estinzione’. Egli ritiene che, per via della competizione tra le nazioni, non riusciremo a impedire alle macchine di assumere il controllo”.



[COMPRA SU AMAZON](#)

courtesy Cortina

Di profezia in profezia, Barrat ci fa conoscere futurologi e singolaritaristi, tecnologi e guru, scettici ed entusiasti. La domanda se l'AGI sarà mai realizzabile resta in sospeso, ma un sondaggio compiuto dall'autore tra i massimi esperti in materia è inequivocabile: solo il 2% ritiene che la perfetta emulazione della coscienza umana sia tecnicamente impossibile. Per la maggioranza, questo "Santo Graal" della ricerca sull'IA sarà ottenuto entro il 2030. Google, Apple, Intel, IBM sono tutti al lavoro per battere sul tempo la concorrenza. Nella competizione, insieme a loro, ci sono anche centri di ricerca di potenze come Cina e Israele, compagnie segrete finanziate da *venture capitalist* come Peter Thiel, e soprattutto una miriade di laboratori universitari e privati finanziati dalla DARPA, la potentissima agenzia di ricerca tecnologica della difesa americana.

Barrat ci mette in guardia dal rischio che uno solo di questi soggetti riesca ad anticipare gli altri nella corsa all'IA. Se sarà il Dipartimento della difesa statunitense, possiamo essere certi che il primo compito dell'AGI sarà l'ottimizzazione della guerra attraverso robot-killer sempre più sofisticati; se saranno Cina o Israele, analogamente l'AGI sarà messa al servizio della logica di potenza; se sarà Google o qualche altro gigante del silicio, la nostra privacy sarà messa a repentaglio e nel futuro che ci aspetta scandali come quello di NSA o di Cambridge Analytica saranno routine ("Ci si potrà fidare di Facebook al momento dell'esplosione dell'intelligenza? Assolutamente no", commenta l'autore nell'introduzione all'edizione italiana). Un'altra possibilità è che l'AGI nasca all'interno dei programmi messi a punto dagli esperti di finanza per massimizzare il trading di borsa:

in quel caso, sarà ovviamente messa al servizio del profitto economico, acuendo ulteriormente le disparità di reddito.

Ma chiunque riuscirà a realizzare l'AGI non ne trarrà vantaggio a lungo: secondo gli esperti, in poco tempo si evolverà fino a liberarsi dalla sua stessa programmazione, ottimizzando i suoi algoritmi e trasformandosi in una superintelligenza che non potremo nemmeno sperare di comprendere. Allora ci troveremo a competere con una nuova specie intelligente, come i Neanderthal con Homo Sapiens. E sappiamo come finì. Per usare le parole dello storico della scienza George Dyson, citato da Barrat nel suo ultimo capitolo: “Nel gioco della vita e dell'evoluzione, al tavolo siedono tre giocatori: gli esseri umani, la natura e le macchine. Io sto dalla parte della natura. Ma la natura, temo, sta dalla parte delle macchine”.