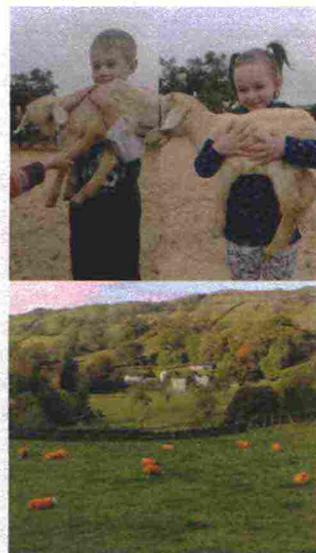


# MA QUANTO È STUPIDA L'INTELLIGENZA ARTIFICIALE

I COMPUTER CI BATTONO SENZA PIETÀ AGLI SCACCHI, MA POI SCAMBIANO UN PRATO PER UNA PECORA (PER NON DIRE DEL LORO RAZZISMO E SESSISMO). DUE RICERCATRICI SPIEGANO PERCHÉ



## ERRORI DI SISTEMA

Sopra, dall'alto, immagini che il software di riconoscimento **Azure** di Microsoft ha valutato erroneamente: fuorviato dai molti esempi visti nella fase di istruzione, ha scambiato le pecore in braccio ai bambini per gatti e considerato fiori gli animali su un prato colorati artificialmente. Sotto, il software **Compas**, utilizzato dal sistema giudiziario americano per valutare i detenuti ai quali concedere la libertà condizionale: si è rivelato razzista

di **Giuliano Aluffi**

**L**A CHIAMANO intelligenza artificiale. Sull'artificiale, niente da dire. Ma sull'intelligenza? Gli algoritmi che si rivelano prodigiosi in ambiti super-specialistici come gli scacchi o il Go (antico gioco da tavolo cinese) sono ancora in grande e sorprendente difficoltà di fronte a compiti in cui eccellerebbe un bambino. Come riconoscere il genere di una persona in una fotografia, operazione che – secondo uno studio di Joy Buolamwini, ricercatrice informatica del Mit – non riesce nel 35 per cento dei casi di fronte a foto di donne di colore, nel 12 per cento dei casi di fronte a uomini sempre di colore e nel 7 per cento dei casi con donne dalla pelle chiara. Solo per i maschi bianchi il riconoscimento del sesso dal volto avviene senza errori.

Joy Buolamwini, afroamericana, ha intrapreso questo studio dopo aver notato come sia al Georgia Institute of Technology che al Mit i software di riconoscimento facciale riuscissero a identificare i suoi amici bianchi ma non lei: veniva riconosciuta come una

donna solo quando si metteva una maschera bianca.

«Il motivo è semplice, l'algoritmo è stato allenato dai suoi programmatori usando archivi fotografici con una presenza straripante di soggetti maschi bianchi e una percentuale irrisoria di persone di colore, tra cui pochissime donne. Per lo stesso motivo gli algoritmi di riconoscimento vocale fanno più errori con le voci femminili, e certi asciugamani elettrici non si attivano quando sotto la loro fotocellula c'è la mano di una persona di colore» spiega Janelle Shane, ricercatrice informatica, blogger (con *AiWeirdness.com* tratta il lato

umoristico dell'intelligenza artificiale) e autrice di *You Look Like a Thing and I Love You* (Headline, pp. 272, euro 24,80). Il titolo del libro – *Sembra una cosa e ti amo* – è stato scelto da un algoritmo progettato da Janelle Shane per generare complimenti affettuosi ma inediti a partire dalle tipiche frasi da latin lover trovate su internet. E tutto il saggio smitizza l'intelligenza artificiale, esplorandone i limiti. La "stupidità artificiale" ha causato infatti nel 2016, in Florida, un incidente mortale a bordo di una Tesla: «Il guidatore stava usan-

**I SOFTWARE ALLENATI CON TROPPI VOLTI MASCHILE BIANCHI, FANNO ERRORI CON DONNE E NERI**

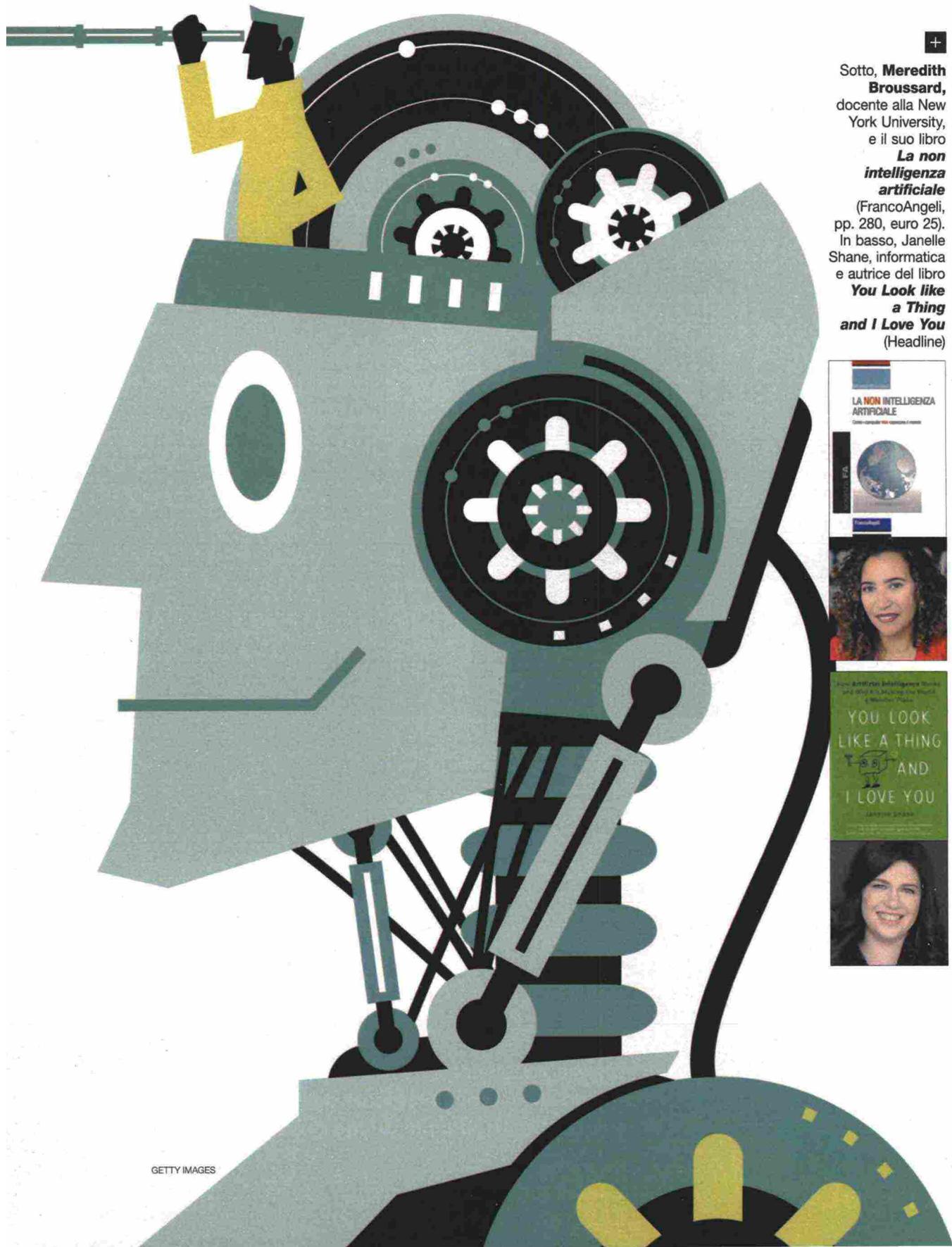
Il software di riconoscimento facciale di Microsoft Azure ha valutato erroneamente le immagini sopra. Sotto, il software Compas, utilizzato dal sistema giudiziario americano per valutare i detenuti ai quali concedere la libertà condizionale: si è rivelato razzista

Two Petty Theft Arrests	
<b>VERNON PRATER</b> Prior Offenses 2 armed robberies, 1 attempted armed robbery Subsequent Offenses 1 grand theft <b>LOW RISK 3</b>	<b>BRISHA BORDEN</b> Prior Offenses 4 juvenile misdemeanors Subsequent Offenses None <b>HIGH RISK 8</b>

Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not confess.

Two Petty Theft Arrests	
<b>VERNON PRATER</b> <b>LOW RISK 3</b>	<b>BRISHA BORDEN</b> <b>HIGH RISK 8</b>

Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not confess.



GETTY IMAGES

+

Sotto, **Meredith Broussard**, docente alla New York University, e il suo libro **La non intelligenza artificiale** (FrancoAngeli, pp. 280, euro 25). In basso, Janelle Shane, informatica e autrice del libro **You Look like a Thing and I Love You** (Headline)



do su strade urbane un software di autopilota che era stato allenato sulle autostrade, dove i camion si vedono solo da dietro e mai di traverso» ricorda Shane. «Così, di fronte a un camion che si era messo di traverso alla strada, l'autopilota lo ha scambiato per un innocuo cartello sopravelevato, e ha deciso di non frenare».

**SCORCIATOIE MALSANE**

Il problema sta nel modo in cui l'intelligenza artificiale apprende: attraverso gli esempi. «Un algoritmo di I.A. ne scorre milioni cercando di scoprire schemi ricorrenti, regole e correlazioni di cui lo stesso programmatore umano può essere all'oscuro» spiega Shane. «Quando abbiamo di fronte un gatto, noi sappiamo all'istante cos'è, ma se dovessimo specificare a un computer qual è la sequenza di regole per riconoscere un gatto, saremmo in difficoltà. Per fortuna, all'intelligenza artificiale non serve ricevere istruzioni esplicite né regole generali: un algoritmo può osservare decine di migliaia di fotografie di gatti e trovare per conto suo le regole che permettono di identificarne uno con successo. Vale ancora di più per giochi come gli scacchi o il Go, dove il numero di mosse possibili è così enorme che nemmeno un maestro saprebbe stilare un insieme di regole per trovare la mossa migliore in ogni situazione. Un algoritmo però può giocare milioni di partite contro sé stesso e derivarne regole e strategie – anche molto inusitate per gli umani – che lo aiutano a vincere nuove partite». I disastri avvengono quando l'insieme di esempi iniziali che l'algoritmo usa per definire le sue regole non è rappresentativo della realtà che il sistema dovrà affrontare, oppure quando l'algoritmo trova "scorciatoie" efficaci ma malsane.

Un esempio curioso è la visione di animali inesistenti. «Un giorno, usan-

do il software di riconoscimento immagini di Microsoft, "Azure", ho notato una bizzarria. L'algoritmo vedeva pecore in fotografie dove c'erano solo prati verdi» dice Shane. «Poi ho capito: l'algoritmo era stato allenato con foto di pecore al pascolo, e non aveva compreso che l'etichetta di "pecora" andava abbinata solo all'animale, e non ai prati erbosi. A riprova: foto di pecore in ambienti inusuali per loro – come un soggiorno – venivano etichettate come "gatti". È impossibile prevedere quali bizzarre scorciatoie un algoritmo potrà escogitare, ecco perché è importante assicurarsi che, perlomeno, l'insieme dei dati iniziali non abbia qualche vizio».

Se ne sono accorti anche gli esperti di intelligenza artificiale di Amazon, che nel 2014 hanno iniziato a lavorare a un algoritmo per selezionare i curriculum più interessanti tra le migliaia che ogni giorno arrivano nelle loro caselle email. «Per allenare il sistema a riconoscere i curriculum "degni di Amazon", sono stati usati i dati sulle assunzioni degli ultimi dieci anni. Ma quei dati storici registravano molti più uomini che donne tra i colloqui

superati con successo, per via della (classica) sovrarappresentazione maschile nelle professioni informatiche» spiega Meredith Broussard,

docente di Data Journalism alla New York University e autrice di *La non intelligenza artificiale. Come i computer non capiscono il mondo* (Franco Angeli). «Quindi alla fine l'algoritmo non cercava effettivamente i candidati migliori, ma quelli che più rispecchiavano le scelte (sessiste) fatte in passato. Si era "convinto" che bisognasse assumere soprattutto maschi, e che pertanto i curriculum delle donne fossero da penalizzare. I programmatori inserirono una correzione: l'algoritmo non doveva usare il nome del candidato come elemento di discriminazione. Ma quello nel frattempo aveva imparato a identificare le donne anche da parole come "squadra di calcio femminile" o "collegio femminile" trovate nei curriculum». Alla fine Amazon accantonò il progetto.

**PREGIUDIZIO RAZZISTA**

Un algoritmo che invece è stato applicato per un anno, producendo vere e proprie discriminazioni razziali scoperte nel 2017 dai giornalisti investigativi del sito ProPublica, è Compas, software di intelligenza artificiale utilizzato come strumento di supporto ai giudici. «L'algoritmo forniva, per ogni detenuto candidato alla libertà condizionale, una stima delle probabilità di recidiva» spiega Broussard. «Ma siccome i dati con cui era stato allenato riflettevano i pregiudizi razziali del sistema giudiziario americano, dove per una persona di colore è più facile finire dietro alle sbarre, l'algoritmo sovrastimava di molto la probabilità di recidiva dei detenuti di colore». Così Compas ha stimato "ad alto rischio di recidiva" Brisha Borden, ragazza di colore che aveva alle spalle tre furti di biciclette commessi da minorenni, mentre ha giudicato "a basso rischio" il bianco Vernon Prater, autore di tre rapine a mano armata. Di fatto dopo il rilascio la Borden non si macchiò più di reati, mentre Prater commise un'altra rapina aggravata. Con un complice molto freddo e determinato ma assai poco intelligente: l'algoritmo Compas.

**Giuliano Aluffi**

**GLI ALGORITMI SONO PRODIGIOSI IN AMBITI SUPER SPECIALISTICI, COME IL GIOCO DA TAVOLO GO**



Una sfida a Go: in giochi come questo o gli scacchi, l'intelligenza artificiale ormai è in grado di battere quella umana