

L'algoritmo tra volontà e rappresentazione

di *Andrea Gatti*

Abstract: The Algorithm between Will and Representation – The present article deals with some theoretical issues concerning human and A.I. bias. Algorithm is supposed to be a means to prevent biased/not objective decisions. Nevertheless, we experience on the daily basis that human interaction (and even “irrationality”) must always occur in determining a judicial decision. This “irrationality” sometimes avoids and sometimes provides justice.

Keywords: Algorithm; Trial; Bias; Discrimination.

1. L'algoritmo come velo di Maya che svela l'irrazionalità delle decisioni umane

3457

Il vasto respiro delle relazioni principali in rassegna rende più difficile proporre delle considerazioni che non appaiano irrilevanti o quantomeno puramente accessorie. Tra i vari temi trattati pare qui opportuno soffermarsi sull'aspetto legato alle criticità dell'algoritmo usato all'interno del processo. Da molte parti ci si chiede come si rapporta l'algoritmo con la creatività dell'interpretazione, con il diritto allo sviluppo del diritto che è implicito in ogni sistema che abbia una sua coerenza intrinseca, ma anche con le clausole generali (ad esempio l'“interesse del minore” ex art. 216 c.c.) o con le formule processuali (ad esempio il modo che il giudice ritiene più opportuno nella conduzione dell'istruttoria, ex art. 702ter c.p.c.). Alla base di tutti questi interrogativi ci sono sempre due questioni fondamentali di ordine teorico-generale: quelle che riguardano il problema dell'indipendenza del giudice (di cui all'art. 101 della Carta costituzionale) e quelle che riguardano il problema del pregiudizio che può scaturire dalle decisioni algoritmiche. Le due questioni sono strettamente legate.

Sia l'intervento di Elettra Stradella che quello di Carlo Casonato hanno evidenziato i *bias* della decisione algoritmica¹. Uno dei concetti che stanno alla base del loro ragionamento (giustamente trattato *en passant* perché considerato ormai come acquisito) è che la logica inferenziale che guida i processi decisionali

¹ Sul rapporto tra pregiudizio del giudice e pregiudizio/errore algoritmico si vedano, non a caso, proprio i recenti contributi di E. Stradella, *Stereotipi e discriminazioni: dall'intelligenza umana all'intelligenza artificiale*, in *Consulta Online*, 20 marzo 2020 (anticipazione del *Liber Amicorum* per Pasquale Costanzo), online su www.giurcost.org, in part. 9 ss. e C. Casonato, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo* - Speciale 2019, 101-130.

(giudiziari, amministrativi e, nel futuro, anche legislativi) si basa su degli *input* che costituiscono la sua base di conoscenza. Pertanto, quando si dice che gli algoritmi (e anche gli algoritmi predittivi) si basano su decisioni già prese, ciò porta a concludere che non è l'algoritmo a contenere il *bias*, ma che ad essere discriminatoria è la sequenza storica delle decisioni che sta a monte. Questo emerge con chiarezza dai casi statunitensi più eclatanti sulla determinazione della probabilità di recidiva e sulla sovrastima del tasso delinquenziale della popolazione di colore rispetto a quella bianca calcolato sulla base dell'algoritmo (si pensi al celebre *State vs Loomis*), ma anche dai recenti problemi che si stanno riscontrando nella censura degli *hate speech* da parte delle piattaforme². In questi casi – come negli altri che sono stati citati soprattutto dalla relazione di Carlo Casonato – l'algoritmo assume anche a strumento per indagare la validità delle forme di legittimazione del giudice. Ciò vale in particolare sotto due aspetti, il primo legato alla natura “interattiva” del processo, il secondo legato alla componente irrazionale che lo abita.

Innanzitutto, anche se è vero che diritto e processo sono sequenze logiche, sono diagrammi di flusso, sono un insieme di istruzioni che consentono di risolvere problemi trasformando i dati iniziali in risultati, è altrettanto vero che la sottoposizione di una controversia a postulati formali o automatici è incompatibile con la realtà giudiziaria. Quest'ultima implica, per sua natura, un contraddittorio dialogico che presuppone un'interazione reticolare tra le parti che è fatto di contestazioni, discussioni, connessioni logiche, in breve di razionalità argomentata³. Durante il processo le variabili iniziali possono cambiare proprio grazie all'interazione umana.

In seconda battuta, la precisione algoritmica (che è anche precisione nell'errare) ci dimostra anche tutto il sottofondo irrazionale che sta alla base delle decisioni giudiziarie “umane”. Sotto questo profilo la dimensione umana del giudice, pur oggetto di accese discussioni, non è di per sé contestata: anzi essa talvolta – oltre che legittima – sembra essere addirittura necessaria: laddove una legge o un orientamento giurisprudenziale siano manifestamente obsoleti (rispetto all'opinione comune) o ingiusti (si pensi al celebre caso della Corte Suprema USA *Lawrence v. Texas* sulla omosessualità come reato) o contrari ad interessi strategici nazionali⁴.

Appare chiaro come qui il vero dilemma teorico che si pone è la generale tolleranza che si percepisce per il condizionamento politico del giudice, ma non per

² J. M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, in *Univ. California L. Rev.*, vol. 51, 2018, 1151-1210. Più risalente S. M. Benjamin, *Algorithms and Speech*, in *Univ. Pennsylvania L. Rev.*, vol. 161, 2013, 1445-1494; per la dottrina italiana G. L. Conti, *Manifestazione del pensiero attraverso la rete e trasformazione della libertà di espressione: c'è ancora da ballare per strada?*, in *Rivista AIC*, n. 4, 2018, 200-225.

³ P. Moro, *Intelligenza artificiale e professioni legali. La questione del metodo*, in *Journal of Ethics and Legal Technologies*, n. 1, 2019, 24 ss.

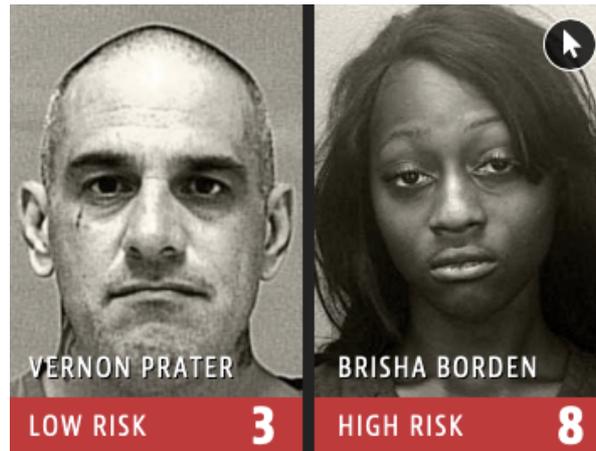
⁴ La salvaguardia di interessi strategici costituisce il tema del ben noto “caso Ilva” che ha portato all'annullamento del sequestro da parte della Corte di Cassazione nonostante le “gravi violazioni ambientali riscontrate”. Cfr. Cassazione, sez. VI penale, 20 dicembre 2013 (depositata 21 gennaio 2014), n. 3635. Anche qui è lecito chiedersi come avrebbe deciso l'algoritmo date le “gravi violazioni ambientali riscontrate” dalla Corte.

quello dell’algoritmo. Il misoneismo che può nascere nei confronti dell’applicazione dell’algoritmo nel processo (ma lo stesso potrebbe valere per la decisione algoritmica amministrativa e per quella legislativa) è dato allora anche dalla sua capacità di svelare la natura convenzionale della legittimazione del giudice (non ci si riferisce ovviamente alla legittimazione formale) e di svelarne la sua dimensione di “umanità”. L’intelligenza artificiale applicata al merito delle decisioni prende quindi le forme di un vero e proprio velo di Maya: mette l’interprete di fronte ad uno specchio mostrando in che misura le decisioni pre-algoritmiche fossero condizionate da molti fattori extra giuridici (suggerimenti politiche, pressioni mediatiche, il *clamor fori*...). L’algoritmo riveste qui la stessa funzione del fanciullo della fiaba di Andersen che addita il re nudo e svela così ciò che tutti pensano, ma nessuno dice: che forse non esiste decisione che sia autenticamente neutrale o imparziale e che ogni scelta è necessariamente condizionata dai pregiudizi (soggettivi) del decisore o, in altri casi, da esigenze “politiche” (oggettive).

2. L’irrazionalità umana come rimedio alla cecità algoritmica

Dal lato opposto, tuttavia, la rivendicazione della “bontà” del pregiudizio umano è anche uno degli strumenti di cui si è serviti per dimostrare l’inadeguatezza dell’algoritmo in alcuni campi. Nell’applicazione della legge – pensiamo all’ambito penale – non c’è solo il mero automatismo induttivo: ci sono dei casi in cui il giudice è chiamato a valutare con una certa discrezionalità la durata della pena (si pensi soprattutto al sistema statunitense) o l’applicazione di una sanzione accessoria o, al contrario, di un beneficio di legge. Esemplicativi sono i due casi statunitensi Vernon Prater e Brisha Borden, richiamati anche da Carlo Casonato nel suo intervento. Prater, un pluripregiudicato per varie rapine a mano armata, era stato giudicato dall’algoritmo un soggetto a basso rischio recidiva. Borden, anche lei pregiudicata ma per un reato minore, era stata giudicata ad alto rischio. Secondo i commentatori il risultato della decisione algoritmica era dipeso proprio da quelle valutazioni circa l’appartenenza “razziale” (la comunità nera statisticamente più esposta alla recidiva di quella bianca)⁵. Una decisione basata su un pregiudizio, ovviamente, ma in questo caso un pregiudizio portato avanti dall’intelligenza artificiale. Per dimostrare la “cantinata” presa dall’algoritmo – per riprendere l’espressione usata da Carlo Casonato nel suo intervento – vengono mostrate le foto segnaletiche dei due imputati.

⁵ C. Castro, *What’s Wrong with Machine Bias*, in *Ergo*, vol. 5, n. 15, 2019-2020; J. Agwin *et al.*, *Machine Bias*, in *ProPublica*, 23 maggio 2016, online su www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.



Qui il giudizio critico nei confronti della predittibilità deriva a sua volta da un altro pregiudizio, vale a dire dall'intuito personale. Avuto riguardo ai rispettivi precedenti penali e ai rispettivi fenotipi (e dunque affidandosi al proprio "pregiudizio", l'intuito), il giudice umano – sembrano suggerire i commentatori – nei casi Prater e Borden avrebbe con tutta probabilità fatto scelte diverse. Così, il pregiudizio umano sarebbe stato in questo caso più "efficace" della conclusione – formalmente corretta ma sostanzialmente ingiusta – raggiunta attraverso il ragionamento algoritmico. Il pregiudizio sarebbe stato infatti "legittimo" e in un certo senso "fondato" perché si sarebbe formato su una certa base di esperienza (anche se pseudo-razionale) e avrebbe mirato al perseguimento della giustizia nel caso di specie⁶.

Da quanto detto si può ricavare una conclusione di massima: che il pregiudizio valutativo tanto umano quanto algoritmico possa superarsi solo laddove la decisione dell'intelligenza artificiale sia supervisionata da quella umana, sulla falsariga dell'apporto del perito nel processo, in cui il giudice rimane il vero titolare del potere decisorio ed egli decide, se del caso, anche contro l'opinione dei suoi periti. È solo da questa *joint venture* che potrebbero bilanciarsi le rispettive incongruenze⁷.

3. La giustizia come rappresentazione

Sia che l'algoritmo prenda la decisione "giusta" sia che prenda quella "sbagliata" (come nei casi Prater e Borden), esso comunque rappresenta un mezzo che ci aiuta a svelare la natura delle decisioni e, in certi casi, a metterci di fronte alla nostra cattiva coscienza come fa, appunto, il velo di Maya. E forse non è inutile fare riferimento al pensiero di Arthur Schopenhauer (che dell'immagine del velo di

⁶ Proprio nel commentare i due casi in oggetto, Carlo Casonato affermava nel suo intervento «la giustizia è altro rispetto ad un giudizio di probabilità statistica».

⁷ Così, ad esempio, I. A. Nicotra, V. Varone, *L'algoritmo, intelligente ma non troppo*, *Rivista AIC*, n. 4, 2019, 87-106: «seppur programmat[o] al massimo grado di precisione, [l'algoritmo] non può mai sostituire interamente l'attività cognitiva, acquisitiva e di giudizio che costituisce estrinsecazione dei valori costituzionali di uguaglianza, correttezza, imparzialità» (90)

Maya ha fatto il marchio della propria filosofia): secondo il grande filosofo tedesco il mondo che ci circonda esiste solo come rappresentazione, vale a dire sempre e soltanto in rapporto ad un “altro” e questo “altro” è colui che lo rappresenta. Se la scienza (intesa come volontà) non è in grado di spiegare tutto, questo lo potrà fare il pensiero umano (inteso come rappresentazione), che è empirico e procede per analogia. Traslando la metafora sul piano del diritto e del processo, ciò non significa altro che il carattere noumenico della giustizia più che nella volontà (l'algoritmo)⁸, sta forse nella rappresentazione della realtà⁹.

Andrea Gatti
Dip.to di Giurisprudenza
Università di Pisa
andrea.gatti@jus.unipi.it

⁸ È anche in questo senso che penso sia interpretabile la felice espressione di A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *Rivista di BioDiritto*, n. 1/2019, secondo cui l'errore algoritmico deriva dal suo «essere il dover essere».

⁹ Schopenhauer stesso ne *Il mondo come volontà e rappresentazione* riconosceva proprio nella giustizia il primo dei due stadi per superare «i mali della volontà».