

SALVATORE AMATO\*

## Emozioni sintetiche e sortilegi al silicio\*\*

### ENGLISH TITLE

Synthetic Emotions and Silicon Predictions

### ABSTRACT

The algorithm is a sublimation of statistics. If statistics is still a reflection of reality, the algorithm becomes reality in itself, or self-reference, or implicit knowledge: a pure abstraction that evolves independently of any human intervention. It is no longer man who knows by algorithms, but the algorithms know by themselves. Artificial intelligence is the reflection of statistics; it is the extreme and definitive passage from the real to the virtual. We have reached the decisive turning from the humanization of the divine, the sovereign as a mortal God, to the deification of computation; a technological, social and cultural revolution called infinite computing. We don't know how many autonomous software agents reshaping our lives, our interactions, and our environments. Faced with a systematic and overall change in human condition, what kind of society can we expect?

### KEYWORDS

Algorithms – Self-Reference – Quantified Self – Black Box World – Meaningful Human Control.

### 1. CALCOLEMUS

*Non disserto, sed computo* è la premessa epistemologica che consentiva ad Hobbes di non nutrire dubbi sulla rigerosità della sua analisi. Si riteneva al riparo da ogni critica, perché la filosofia vera «(cioè esatta)» non si fonda sul «belletto del discorso», ma sul ragionamento che «è la stessa cosa che addizionare e sottrarre»<sup>1</sup>. *Calculemus* invocava Leibniz, per sedare le dispute

\* Professore ordinario di Filosofia del diritto presso il Dipartimento di Giurisprudenza dell'Università di Catania.

\*\* Prodotto del Programma di ricerca di Ateneo UNICT 2020-22 linea 2.

1. Hobbes (1655), trad. it. 1972, pp. 70-71.

tra i filosofi<sup>2</sup> e per propugnare una razionalizzazione dell'azione politica. Foucault ha definito *mathesis* questo spirito che, a partire dal XVII secolo, attraversa la nostra cultura creando una singolare connessione tra algebra e tassonomia: entrambe sono scienze dei segni, l'una si applica alle nature semplici e l'altra alle nature complesse; entrambe sovrappongono un ordine fittizio all'ordine reale, sostituendo simboli e calchi ai rapporti tra le cose. L'analisi di Foucault ha, come sempre, una straordinaria efficacia evocativa nel descrivere come «nell'ambito di un siffatto sapere, si tratta di apporre un segno a tutto ciò che può offrirci la nostra rappresentazione: percezioni, pensieri, desideri; questi segni debbono valere in quanto caratteri, cioè articolare l'insieme della rappresentazione in aree distinte, separare le une dalle altre mediante tratti assegnabili; essi autorizzano in tal modo l'instaurarsi di un sistema simultaneo in base al quale le rappresentazioni enunciano la loro prossimità e la loro lontananza, la loro continuità e i loro scarti – quindi il reticolo, acronologico, che ne manifesta la parentela e ne restituisce in uno spazio permanente le relazioni d'ordine»<sup>3</sup>.

Sarà l'ingresso ufficiale della statistica tra le materie di insegnamento a dare a questa connessione tra algebra e tassonomia una specifica connotazione scientifica e un definito carattere politico come studio analitico e descrizione sistematica delle caratteristiche e delle potenzialità dello Stato. La statistica nasce timidamente come *Staatenkunde* nelle lezioni di Hermann Conring a metà Seicento ed inizia ad assumere, oltre all'attuale denominazione, anche il ruolo di metodo d'indagine dei fenomeni collettivi nell'autorevole insegnamento di Gottfried Achenwall a Göttingen: una prospettiva particolare con cui sviluppare la tradizione del diritto naturale, di cui Achenwall era uno dei maggiori esponenti<sup>4</sup>, collegandola con gli influssi dell'empirismo inglese particolarmente sentiti a Göttingen per effetto dei legami della Bassa Sassonia con la dinastia degli Hannover. Dobbiamo tenere presente che il diritto naturale era, a quel tempo, tanto un elemento essenziale della preparazione dei funzionari statali, quanto il sostegno dell'assolutismo illuminato. Il diritto del sovrano a rappresentare tutti, a decidere su tutto e a preservare il tutto, anche a costo del sacrificio dei singoli cittadini (*homo perducitur in casum necessitatis*), trovava nella pretesa capacità della statistica di penetrare la natura delle cose lo strumento oggettivo con cui individuare e giustificare qualsiasi azione di governo.

La statistica come espressione della scienza dello Stato (*Staatwissenschaft*), come procedura dell'organizzazione del potere (*Staatsverfassung*), ma soprat-

2. Geymonat, Giorello, 1977, p. 380.

3. Foucault, 1967, p. 89.

4. Lo *Ius naturae* di Achenwall, ad esempio, è alla base delle lezioni di Kant sul diritto naturale (*Naturrecht Feyerabend*) recentemente tradotte in italiano (2016).

tutto come scienza della costrizione e del controllo in cui è così e non può che essere così, perché non è la volontà del sovrano a decidere, ma sono i numeri e i conti. Attraverso questa declinazione assolutistica del *non disserto, sed computo* inizia quel processo di «desostanzializzazione»<sup>5</sup> in cui lo Stato tende a presentarsi come un apparato neutrale di assimilazione e di condizionamento delle aspettative e dei comportamenti individuali, una volta che siano standardizzati e omologati per essere adattati alle caselle delle rilevazioni matematiche. Mi pare estremamente significativo che sia stato uno studioso di Machiavelli, Francesco Sansovino, nel *Governo et amministrazione de diversi regni* pubblicato a Venezia nel 1562, a sottolineare come l'esercizio del potere non potesse fare a meno della rigorosità dei calcoli e che sia stato uno studioso della ragion di Stato, Giovanni Botero, a riproporre, una trentina di anni dopo, questo tema nelle *Relazioni universali*.

Se lo Stato si desostanzializza, non perde, per questo, la propria efficacia condizionante né riduce le proprie pretese di controllo e omologazione. Anzi le accresce. Soltanto se i fatti diventano artefatti, la comunicazione informazione, le aspirazioni dati, i dati sequenza numerica, l'occhio vigile del potere può inglobare ogni cosa. La statistica rende tutto visibile e misurabile. Se tutto è visibile, tutto è controllabile. Se tutto è misurabile, tutto è assimilabile. «Nessuna soluzione per nessun problema potrà mai avere la pretesa di essere unica. Se ciò avviene vuol dire che si intende instaurare un nuovo dominio»<sup>6</sup>. Quando lo Stato si appropria della statistica, la sua capacità di controllo diviene totale: è onnipresente e intanto assente nell'autosufficienza dei modelli astratti a cui riconduce ogni cosa svuotata delle proprie peculiarità e avulsa dalla propria identità.

Solo chi ha la capacità di raccogliere e disporre dei dati relativi ai fenomeni collettivi può pretendere di trarre da essi inferenze attendibili sulle quali fondare i processi decisionali e giustificare le prescrizioni imposte alle condotte individuali: la moneta, il fisco, il prezzo delle merci, le leve militari... l'economia nel suo complesso sono, anche, una scienza dei numeri che fonda il benessere del paese sulla capacità di azzeccare le previsioni, sviluppando adeguate prospezioni matematiche.

È questa soggezione al calcolo che ha sempre sorretto il potere nella sua pretesa di guidare, attraverso il controllo del tempo e dello spazio, le condotte umane. Già «nei catasti neo-assiri venivano elencati i bambini misurati a palmi per stimare il numero di contadini adulti disponibili entro un certo numero di anni»<sup>7</sup>. Una delle prime divinità egiziane è Seschet che custodisce i

5. Gentile, 2003, p. 37.

6. Mezza, 2018, p. 33

7. Valerio, 2020, p. 53. Ho, tuttavia, molti dubbi sulla tesi di fondo del libro secondo cui la matematica sia una disciplina che favorisce la diffusione della democrazia. Può farlo, come qualsiasi altra scienza, ma può avvenire anche l'opposto.

numeri a tutela dei tracciati dei campi e della misurazione dei confini. Dagli imperatori cinesi, «padroni unici del calendario e, a tale titolo, animatori di tutta la terra», alla centralità dei *mensores* nella burocrazia romana, chiunque «dio, eroe o capo, voglia trionfare, regnare, fondare: chiunque egli sia, deve tentare di impadronirsi del tempo, allo stesso titolo dello spazio»<sup>8</sup>. C'è sempre qualcosa di rassicurante e insieme di misterioso nella pretesa di possedere la «misura» del futuro perché il controllo del futuro implica il trionfo dell'immateriale sul reale, dell'arcano sul visibile. Cos'è il «sortilegio» se non il condizionamento stringente di una logica sfuggente?

In apparenza ci troviamo di fronte a un fenomeno essenzialmente quantitativo e quindi tendenzialmente avalutativo, ma il margine discrezionale è sempre presente dietro ogni atto di rilevazione, classificazione e organizzazione. È neutrale l'inferenza, ma non l'assunzione dei dati e l'organizzazione dei processi inferenziali. Solo che una volta sviluppata la *mathesis*, una volta trasformata la realtà in segno e assunta la concatenazione dei segni a dimensione della realtà, non è più possibile venirne fuori, non è più possibile falsificarla, perché ogni struttura inferenziale è autoreferenziale. L'autoreferenzialità rende imperscrutabile, e quindi indiscutibile, il fondamento del processo decisionale. Come ci ha spiegato Luhmann l'autoreferenza funziona «come un blocco alla riflessione: ha la funzione di rendere invisibile il decisore della decisione»<sup>9</sup>. È così perché è così. Pensiamo al ruolo centrale, e in qualche modo misterioso, della *blockchain* a fondamento attualmente di numerose catene seriali di documenti che alimentano e certificano lo scambio di decisioni.

Non è un caso se la teoria dell'assolutismo nasce assieme alla desostanzializzazione dello Stato. L'autoreferenzialità determina l'assolutizzazione perché si sviluppa da sé, si astrae da qualsiasi contingenza e si mette al riparo da qualsiasi contingenza. Per comprendere il processo di desostanzializzazione dello Stato non dobbiamo solo pensare alla prospettiva con cui Hobbes trasferisce la concezione cartesiana dell'uomo come macchina animata allo Stato come corpo artificiale<sup>10</sup>. Il passaggio ulteriore, e decisivo, lo mette in luce La Mettrie: il corpo umano è una macchina certo, ma «è una macchina che carica da sé i suoi meccanismi, immagine vivente del moto perpetuo»<sup>11</sup>.

Affermare che il sovrano è un «Dio mortale» sviluppa un modello di autosufficienza riflessiva che cancella qualsiasi vincolo e limite esterno: il potere deriva dal potere come il numero deriva dal numero. Un cuore che regola senza essere regolato, diceva Marsilio<sup>12</sup>. La statistica costituisce la proiezione

8. Dumézil, 1935-36, p. 240.

9. Luhmann, 2013, p. 62.

10. Schmitt (1936-37), trad. it. 1986, p. 83.

11. La Mettrie (1747), trad. it. 1978, p. 182.

12. Marsilio da Padova (1324), trad. it. 1960, p. 200.

tecnica di questa possibilità di esercitare il potere con il calcolo, riducendo la società a quello stesso insieme di linee e punti che la visione meccanicistica della natura delineava per i processi fisici. Il calcolatore prende il posto del calcolante, legando indissolubilmente l'uomo e la macchina. I grandi matematici, Pascal, de Colmar, Babbage, Lovelace, Turing sono ideatori di macchine per calcolare. E le macchine per calcolare sono tali perché consentono di fare a meno dei loro ideatori.

## 2. IL PARADISO È UN COMPUTER POTENTE

Non c'è calcolo che non ne possa implicare un altro e poi ancora un altro. Questa infinita potenzialità autoreferenziale giustifica, allo stesso tempo, la fiducia dell'uomo nella macchina e l'autonomia della macchina dall'uomo, ispirando tanto l'*homme machine* di La Mettrie<sup>13</sup> quanto l'addizionatrice meccanica di Pascal fino alla macchina computazionale di Turing e poi ... arriviamo alla svolta della seconda metà del secolo scorso in cui incominciano ad apparire le prime *calculating machine* seguite dalle *computing machine* e dalle attuali *learning machine*. Sempre negli stessi anni si andava affermando l'impero economico dell'IBM (*International Business Machine*) che già nel nome annunciava la nuova era globale delle macchine commerciali, segno emblematico dell'affermarsi del modello di «tecno-scienza-economia» in cui sono le esigenze del mercato a guidare i processi di innovazione e trasformazione. Oggi le c.d. *big five* o GAFAM (Google, Apple, Facebook, Amazon, Microsoft) hanno fatto dell'acquisizione e manipolazione dei dati la maggior fonte di profitti e di potere. La continuità tra Stato e mercato riposa sulla *data sovereignty* in cui l'economia digitale sposta l'accento dal cittadino al consumatore, dall'interesse pubblico ai vantaggi privati senza alterare il dominio della statistica.

Ormai è impossibile distinguere dove finisca la scienza del calcolo e dove cominci la mera tecnologia sociale. L'una incontra l'altra nell'intelligenza artificiale. Un termine vago che abbraccia un'estrema pluralità di agenti, strumenti e funzioni in continua evoluzione e su cui si concentrano gli sviluppi dei vari settori delle tecnologie avanzate. Non c'è forma di sapere che non sia coinvolta in progetti di intelligenza artificiale e non ci sono sviluppi dell'intelligenza artificiale che non coinvolgano tutte le forme di sapere. Il *calculemus* di Leibniz è ormai il fulcro di una società integralmente pervasa di matematica: «algoritmi, previsioni, automazioni, calcoli, cronometri, GPS, conteggi energetici per perdere peso o acquistarne, lotterie, contapassi»<sup>14</sup>.

13. Anche La Mettrie (1978, p. 358), come Hobbes, sostiene di limitarsi alla neutralità della descrizione: «... non moralizzo, non predico, non declamo: spiego».

14. Valerio, 2020, p. 11.

Non si tratta solo di rendere «sensibile», vale a dire in grado di assimilare, memorizzare e trasmettere informazioni, tutto quello che ci circonda, ma di collegarlo in *cloud* ampliando sempre più gli orizzonti dell'infosfera, cioè di tutti quei dati veicolati, conservati e rielaborati dall'insieme delle tecnologie della comunicazione e dell'informazione (*Information and Communication Technologies* – ICT). Come sottolinea Floridi, «la modernità è diventata ben presto un universo di complesse dipendenze reticolari, di reazioni a catena meccaniche così come di connessioni obbligate...»<sup>15</sup>. Un sistema di autoreferenze. L'autoreferenza del calcolo si trasferisce all'autoreferenzialità delle macchine. È ancora Floridi a ricordarci che per la prima volta nella nostra storia disponiamo di tecnologie che possono agire regolarmente e normalmente come utenti autonomi di altre tecnologie<sup>16</sup>.

Pare siano otto miliardi e seicento milioni gli oggetti connessi al *web* che ci assistono nelle nostre più svariate attività. Treni e linee ferroviarie, aerei e rotte aeree, denaro e transazioni finanziarie, messaggi e sistemi di comunicazione, sono regolati da un complesso insieme di astrazioni matematiche fondate sulla statistica. L'intelligenza artificiale è l'implicito meccanismo regolatore della rete neurale (*neural network*) che tiene assieme i processi di connessione tra algoritmi che apprendono da algoritmi (*machine learning*) in successivi strati di stringhe logiche sempre più complesse (*deep learning*). «Una specie di alchimia... che trasforma dati e predizione in miliardi di dollari, e ciò offre una grande sfera di cristallo che contiene e abbraccia qualsiasi ambito della quotidianità»<sup>17</sup>.

Anche se non ce ne rendiamo conto, l'intelligenza artificiale è la destinazione continua delle nostre richieste e intanto l'artefice di un mondo a sé che si accresce continuamente a un ritmo vertiginoso, trasformando il reale in virtuale. È stato calcolato che ogni minuto del 2018 Google ha effettuato 3,88 milioni di ricerche, e le persone hanno guardato 4,33 milioni di video su YouTube, inviato 159.362.760 mail, pubblicato 473.000 tweet, e 49.000 foto su Instagram. Un calderone di dati in cui tutto si accumula senza limiti in un disordine solo apparente, perché ogni *input* e ogni *output* avviene sotto il controllo di specifici processi matematici connessi ad altri processori matematici e poi ancora ad altri. «Alcuni di questi dati vengono usati per migliorare prodotti o servizi, ma il resto diviene un surplus comportamentale privato, sottoposto a un processo di lavorazione avanzato noto come «intelligenza artificiale» per essere trasformato in prodotti predittivi in grado di vaticinare cosa faremo immediatamente, tra poco e tra molto tempo. Infine, questi prodotti predittivi vengono scambiati in un nuovo tipo di mercato per le previsioni

15. Floridi, 2017, p. 31.

16. Ivi, p. 36.

17. Vespignani, Rijntano, 2019, p. 74.

comportamentali, che io chiamo mercato dei comportamenti futuri. Grazie a tale commercio i capitalisti della sorveglianza si sono arricchiti straordinariamente, dato che sono molte le aziende bisognose di conoscere i nostri comportamenti futuri»<sup>18</sup>.

Un mondo di statistiche: l'integrale assimilazione tecnologica di algebra e tassonomia in una *mathesis* che non condiziona solo i modelli di sapere, ma anche i modelli di vita. Si appone un segno, e quindi un algoritmo, a tutto ciò che indica un frammento delle azioni o dell'identità umana, mettendo assieme aree distinte e distanti, come i mille passi che ho accumulato ieri o le parole che digito oggi, unite nella dimensione senza tempo e senza spazio di una stringa numerica. Quel reticolo acronologico di relazioni d'ordine, di cui ci parla Foucault; un reticolo che instaura un sistema simultaneo di rappresentazioni che annullano prossimità e lontananza, continuità e fratture, creando uno schema permanente che contiene tutto e che si propone di gestire tutto come il sovrano illuminato del XVIII secolo. Solo che il posto del sovrano oggi è preso da una pluralità di agenti privati che si sovrappongono, a vari livelli, ai tradizionali operatori pubblici, rendendo ancor più evidente la desostanzializzazione dello Stato, ma non l'effettività e pervasività del potere.

Un insieme di dati non è un'informazione, come un insieme di numeri non è un algoritmo, come un insieme di pietre non è una piramide. È la scienza statistica che consente di tradurre una massa amorfa di elementi in un insieme di istruzioni con cui analizzare la realtà e sviluppare previsioni. La statistica è una stringa numerica orientata operativamente che rappresenta, nel passaggio dalle condotte ai numeri, ciò che esiste mediante ciò che non esiste. L'algoritmo è quella stringa numerica che consente di schematizzare questo passaggio e riprodurlo all'inverso, dai numeri alle condotte, dando consistenza a un'ipotesi. Lo studio della statistica, quando nasce nell'università di Gottinga ai primi a metà del Settecento, affida agli esseri umani il compito di elaborare dati e formulare previsioni. Ora gli algoritmi cedono alle macchine, ai sistemi informatici, questo compito: «possono essere considerati degli invisibili, ma dinamici mediatori logici che agiscono con elevata competenza nei meccanismi degli innumerevoli dispositivi digitali che ci circondano»<sup>19</sup>.

L'algoritmo è la sublimazione della statistica. Quello che nella statistica è ancora un riflesso della realtà, nell'algoritmo diviene la realtà in sé, o autoreferenza o «conoscenza implicita» che dir si voglia. Un'astrazione pura che tende a svilupparsi indipendentemente da ogni intervento umano. Il suo esito estremo potrebbe essere la singolarità, quella situazione in cui la tecnologia si espande a una velocità non più controllabile, per cui i computer saranno

18. Zuboff, 2019, p. 19.

19. Peres, 2020, p. 6.

sempre più intelligenti fino ad arrivare a modificarsi e programarsi in assoluta autonomia, superando l'intelligenza umana<sup>20</sup>.

In queste prospettive non è più l'uomo a conoscere attraverso gli algoritmi, sono sempre più spesso gli algoritmi a conoscere attraverso se stessi. Mi riferisco, indipendentemente da un'ipotetica singolarità, agli attuali algoritmi stocastici o non deterministici che, a fronte degli stessi dati d'ingresso, ammettono più soluzioni tutte ugualmente possibili e tutte ugualmente imprevedibili. Il *black box effect* è la frequente situazione in cui è impossibile da parte degli stessi programmatori e tecnici spiegare come il sistema sia arrivato a determinati risultati e prevedere i risultati stessi. Una previsione non prevedibile non è la fine della statistica, ma il suo definitivo trionfo, perché esalta l'autoreferenzialità: un mondo di statistiche che generano statistiche indipendentemente da qualsiasi controllo umano (*black box world*).

Virtuale deriva da virtù. La virtù non è solo una qualità dell'essere umano, ma esprime l'elemento migliore della sua identità, come indica il legame etimologico con *vir*. *Virtus* era la dea romana del coraggio e del valore militare. È solo nel latino medievale che compare l'aggettivo *virtualis* per connotare qualcosa come possibile, come potenziale in contrapposizione all'effettività materiale del *realis*. Anche *realis* è un costrutto con cui il latino medievale designa, in contrapposizione a *virtualis*, tutto quello che è materiale e concreto. Non vi è ancora il radicale dualismo cartesiano tra *res cogitans* e *res extensa* perché i due mondi del reale e del virtuale si legano assieme: il possibile completa il reale e il reale definisce il possibile. La crescita della tecnologia digitale ci ha costretti a sostantivare l'aggettivo, perché ha rotto questo rapporto, ponendoci di fronte tanto a simulazioni che si sovrappongono alla realtà (l'identità digitale) quanto a simulazioni che si aggiungono alla realtà inglobandola (la memoria digitale). Il virtuale indica un mondo a sé oltre o, forse meglio, altro rispetto alla realtà. A tal punto «altro» che mette definitivamente da parte ogni legame con *vir* per sviluppare una propria «coscienza» al silicio. Alla luce dell'intelligenza artificiale è il reale che esiste in funzione del virtuale. Il paradiso? È «un computer potente»<sup>21</sup>.

C'è sempre l'algoritmo di un algoritmo che rinvia ad altri algoritmi: «estrazione e renderizzazione ubiqua, attivazione (*tuning*, *herding*, condizionamento), catene di rifornimento del surplus comportamentale, processi industriali basati sull'intelligenza delle macchine, fabbricazione di prodotti predittivi, mercati dinamici dei comportamenti futuri e targettizzazione; tutto ciò porta a ulteriori pratiche di *tuning*, *herding* e condizionamento...»<sup>22</sup>. Il nuovo orizzonte che si apre dinanzi ai nostri occhi è la scienza della complessità. Non

20. Kurzweil, 2008, p. 24.

21. Seung, 2013, p. 323.

22. Zuboff, 2019, p. 461.



sappiamo come, non sappiamo perché, non sappiamo neppure se vi sia un «Chi», con la lettera maiuscola, che programma e controlla tutto questo accanto ai tanti «chi», con la lettera minuscola, che regolano il *mini-targeting* sistematico dei nostri comportamenti: biometria, antropometria, psicomelia, fisiognomica, indice cefalico, dinamiche del contagio... Con la miniaturizzazione dei processi di calcolo consentita dagli *smartphone* c'è ormai un algoritmo per ogni aspetto dell'esistenza. Non esiste più un luogo privilegiato di osservazione come poteva essere la demografia che sorreggeva la scienza statistica nel XVIII secolo. Oggi sono infiniti i punti di osservazione, tutti apparentemente marginali e tutti indirettamente rilevanti. «Il panottico digitale del XXI secolo è aprospettico, nella misura in cui non è sorvegliato da un centro unico, dall'onnipotenza dello sguardo dispotico. La distinzione tra centro e periferia, che è costitutiva per il panottico di Bentham, scompare completamente. Il panottico digitale sussiste senza ottica prospettica. Questo ne determina l'efficienza»<sup>23</sup>.

La desostanzializzazione dello Stato rende, inoltre, il potere disponibile a chiunque sia in grado di costruire statistiche alternative, intercettando e condizionando nuove aspettative o alimentando nuovi bisogni. Inviare un *tweet*, apporre un *like*, controllare le pulsazioni cardiache sono diventati le esigenze insopprimibili di un nuovo modello antropologico. Si parla, infatti, di nativi digitali per indicare una generazione la cui dimensione centrale dell'esistenza è la connessione *on line*; una generazione che tende a sostituire i legami virtuali con i legami sociali, il web con lo Stato, il chattare con il parlare. I nativi digitali inviano ogni anno non meno di 350.000 mail o sms, trascorrono 10.000 ore al cellulare, 3.500 sui social network, 5.000 con videogiochi. Il rapporto con il virtuale è indissolubile: ciascuno si immerge nelle tracce proprie e in quelle degli altri, costruisce su queste tracce percorsi, incontri, progetti e addirittura terapie.

È il dominio della statistica. Una statistica onnivale, onnipresente, onnivora, che non trascura neppure il minimo dettaglio e lo ridefinisce secondo schemi e modelli propri, tendenzialmente imperscrutabili. Noi viviamo non solo nella biosfera, da cui traiamo le risorse materiali, ma nell'infosfera da cui deriviamo quelle risorse immateriali che tengono assieme il file che sto scrivendo, il cellulare che sta squillando, Google, Facebook e i tanti bot, robot o computer che utilizziamo. «Ogni giorno, solo nel tempo libero, processiamo 34 gigabyte, equivalenti a 100.000 parole (*Guerra e pace* di Tolstoj è composto da 460.000 parole). Questa cifra ragguardevole ma finita d'informazioni è il frutto di una selezione che operiamo sul totale di oltre 360 esabyte d'informazioni prodotte annualmente dall'uomo, pari a 360 000 000 000 000 000 000 byte. L'accumulo d'informazione è davvero stellare: nel 2014 l'umanità ha

23. Byung-Chul, 2014, p. 62.

generato ogni due giorni una quantità di dati superiore a quella complessivamente prodotta da *Homo sapiens* fino a tutto il 2003»<sup>24</sup>.

Più crescono i dati e più il sistema diviene complesso; più il sistema diviene complesso e più la gestione del sistema si complessifica; più si complessifica la gestione del sistema e più l'intelligenza artificiale tende a divenire generale (un «Chi» con la lettera maiuscola), tende a passare dalla singola applicazione che reagisce ai nostri impulsi a una ridefinizione complessiva che si sviluppa attraverso e oltre i nostri impulsi. L'intelligenza artificiale, che tende a prendere il posto dell'intelligenza umana, è il riflesso della statistica; è l'estrema e definitiva riduzione del reale al virtuale. Dall'umanizzazione del divino, il sovrano come Dio mortale caro ad Hobbes, siamo arrivati alla divinizzazione della computazione, una rivoluzione tecnologica, sociale e culturale «called infinite computing»<sup>25</sup>.

### 3. GLI ALGORITMI COGNITIVI

Come interpretare la prospettiva «infinita» che la computazione sembra aprirci? Per certi versi, sembra allontanarsi da qualsiasi dimensione umana. Le ipotesi estreme della singolarità o del default tecnologico accentuano il rischio di una alterità totalmente estranea se non ostile o addirittura malevola. Anche senza ipotizzare scenari così radicali, ci troviamo già di fronte ad almeno tre tipi di sviluppo: *a*) le decisioni basate-su-algoritmi (*algorithm-based decisions*): sono decisioni integralmente umane, anche se derivano in tutto o in parte da informazioni ottenute mediante calcoli algoritmici; *b*) le decisioni guidate-da-algoritmi (*algorithm-driven decisions*): queste decisioni sono ancora prevalentemente umane, ma in termini molto ristretti rispetto a quanto determinato dall'esito delle procedure informatiche; *c*) le decisioni determinate-da-algoritmi (*algorithm-determined decisions*): sono gli sviluppi dei processi digitali a pervenire alle decisioni, senza alcuna interferenza umana<sup>26</sup>.

Questi diversi algoritmi si riflettono sui modelli di controllo che siamo in grado di esercitare

Dal punto di vista tecnico si delineano quattro prospettive<sup>27</sup>: 1. *human in the loop*: tutte le operazioni avvengono sotto l'integrale controllo umano; 2. *human on the loop*: le macchine agiscono autonomamente, ma rimane una supervisione umana; 3. *human post loop*: gli esseri umani non possono condizionare direttamente l'operato delle macchine, ma possono intervenire *ex post* per bloccare l'azione; 4. *human out of the loop*: le macchine operano in asso-

24. Vineis, Carra, Cingolani, 2020, p. 74.

25. Bass, 2012.

26. *Data Ethics Commission*, 2020, p. 160.

27. Durante, 2019, p. 93

luta autonomia senza nessuna possibilità di controllo umano né preventiva né successiva. È l'ipotesi estrema che pone tanto l'ebbrezza della singolarità quanto il timore di un *default*.

Dobbiamo poi tener presente che non è sempre netta la linea di demarcazione tra gli algoritmi deterministici, che sono integralmente sotto il controllo umano, e gli algoritmi indeterministici o stocastici, che si sviluppano in maniera casuale, perché il passaggio dagli uni agli altri è spesso segnato dal *black box effect*, esiti del calcolo che sfuggono alle nostre previsioni e alla nostra comprensione. Resta alla storia il famoso «colpo 37» con cui il computer *AlphaGo* ha sconfitto il campione del mondo, Lee Sedol, in una partita di *Go*, il gioco più difficile per la sua complessità da inserire in un sistema informatico. Una mossa apparentemente sbagliata, anzi assurda per tutti gli esperti, eppure decisiva per la vittoria. Come può un algoritmo, scritto da esseri umani, fare quello che gli esseri umani non sono in grado concepire? È chiaro che ci troviamo di fronte a un calcolo che ha prodotto un altro calcolo, ma è questo «produrre» che ormai ci sfugge, non siamo più noi i padroni del «sortilegio», non siamo più noi ad appropriarci del futuro.

Ne siamo talmente consapevoli che, attraverso tecnologie estremamente diversificate ed estremamente sofisticate<sup>28</sup>, stiamo provando a elaborare algoritmi generali che possano apprendere rappresentazioni astratte senza che sia necessario l'intervento umano che ne preordini o condizioni gli sviluppi. Insomma dovrebbe venir meno una delle barriere che dividono le macchine dagli esseri umani, quella tra la semantica, l'astrazione simbolica del pensiero che è variabile, creativa e non riconducibile entro una stringa matematica (almeno così credevamo), e la sintassi, la mera connessione dei vari segni aggregabile sistematicamente attraverso costanti matematiche. La differenza tra *dissertatio* e *computatio*, cara ad Hobbes?

L'*Allen Institute for Artificial Intelligence* ha affermato, qualche mese fa, di aver elaborato algoritmi «astrattivi»<sup>29</sup>, capaci di articolare nuove frasi a partire dall'analisi del contenuto dell'intero articolo, mentre finora si erano sperimentati dei sistemi cosiddetti «estrattivi», in grado di individuare solo determinate sequenze linguistiche all'interno del medesimo contesto. Sembra delinearsi una nuova prospettiva sulla riproducibilità della complessità topologica delle reti neurali. Una prospettiva misurabile in termini di «persistenza neurale» in base alla capacità di percepire e rielaborare dati sempre più impliciti e sottintesi. Non sono un informatico e non so dove ci possa condurre tutto questo, mentre l'articolo che cito è pieno proprio di cifre e grafici. Mi limito a

28. Ad esempio macchine neurali, apprendimento profondo per rinforzo, ottimizzazione bayesiana, elaborazione LSTM, word embedding, autoencoder variazionale, generative adversarial network.

29. Cachola, Kyle, Cohan, Weld, 2020.

un'osservazione ingenua. Se la "profondità", come capacità di percepire il «senso» (e quindi di *com-prendere*), opposta alla superficialità, come capacità di individuare i «segni» (e quindi di *ri-prendere*), è uno dei presupposti dell'elaborazione intellettuale, cosa impedisce di ipotizzare che incominci a vacillare anche la barriera tra semantica e sintassi, se non addirittura, tra pensiero e calcolo?

Mi pare estremamente significativa anche la recente esperienza delineata nelle ricerche del *Los Alamos National Laboratory*<sup>30</sup> sulle reti neurali. Anche qui abbiamo modelli di intelligenza artificiale che tendono ad elaborare processi logici sempre più complessi, come lo è la semantica rispetto alla sintassi. Indagano, infatti, sul modo in cui queste reti simulate sono in grado di sviluppare ricerche non supervisionate sui dizionari, classificando gli oggetti senza avere esempi precedenti con cui confrontarli. Lo scopo è quello di sviluppare reti neurali che si avvicinano al modo in cui gli esseri umani e altri sistemi biologici imparano a vedere e a distinguere. L'elaboratore non riproduce nella ricerca le distinzioni con cui è stato programmato, ma costruisce da sé i modelli di classificazione e ricerca. Sempre più semantica e sempre meno sintassi? La cosa straordinaria è che il sistema sembra «avvertire» la difficoltà delle operazioni che compie. Sottoposte ad attività intensa, queste reti mostrano segni di instabilità che vengono meno quando si applicano stati di sonno artificiali, stati analoghi alle onde che i cervelli viventi sperimentano durante il sonno. Sembra, così, possibile resettare i neuroni inattivi per assicurarsi che tornino ad essere componenti funzionali della rete. Questo tipo di instabilità, affermano i ricercatori del *Los Alamos National Laboratory*, non è una caratteristica di tutte le reti di intelligenza artificiale. Il problema sorge solo quando si addestrano processori biologicamente realistici, quando si cerca cioè di riprodurre la stessa funzionalità biologica delle elaborazioni mentali umane. Possiamo paragonare l'instabilità allo stress? L'articolo parla addirittura di sonno e di allucinazioni. Lo stress avvicina gli algoritmi al pensiero?

La costruzione di macchine che funzionano attraverso algoritmi non deterministici presuppone l'elaborazione di una sorta di intelligenza cognitiva, capace di selezionare gli eventi e di adattarsi autonomamente ad essi. Un'intelligenza cognitiva non è ancora un'intelligenza generale in grado, in quanto tale, di elaborare da sé un sistema di regole con cui decidere la propria condotta e predisporre i propri fini, ma, come abbiamo visto negli esperimenti precedenti, è la premessa per la sua realizzazione, perché è estremamente sottile, oltre che tendenzialmente inevitabile, il passaggio che separa un'intelligenza «pratica» programmata sulla base di algoritmi non deterministici per operare tecnicamente allo scopo di raggiungere determinate finalità, da un'intelligenza «sociale» o «ambientale» che adatta autonomamente questi fini al

30. Kenyon, 2020.

contesto in cui opera. Di adattamento in adattamento, da «colpo 37» a «colpo 37», anche l'intelligenza sociale scivola, per svolgere a sua volta con efficienza ogni funzione, verso forme sempre più complesse di intelligenza «teorica» nel senso che si struttura attraverso un sistema esclusivamente «proprio» di regolamentazione.

Tanto è vero che ci nascondiamo dietro il neologismo «agentività», ispirato all'*agency* della terminologia anglosassone ma meno compromettente del nostro «soggettività», per sottolineare questo passaggio dall'intelligenza artificiale «cosa» all'intelligenza artificiale «persona»<sup>31</sup>. Non possiamo escludere che le nuove tecnologie si stiano consolidando attorno a una particolare forma di intelligenza che nasce dalle macchine e si sviluppa con le macchine, andando oltre le macchine. Non sappiamo quanto oltre, ma non possiamo neppure continuare a sostenere che sono solo «una delle tante forme di oggettivazione della soggettività», al massimo degli «oggetti tecnici avvolti in un "alone di socialità" ...»<sup>32</sup>.

Un articolo di *Nature* dello scorso ottobre tocca, già nel titolo, il tema della coscienza umana: *Can lab-grown brains become conscious?*<sup>33</sup>. Siamo in grado di ottenere in vitro strutture miniaturizzate di diversi organi. Tra questi organi anche mini-cervelli cresciuti in provetta che, a quanto pare, sono stati in grado di connettersi autonomamente al midollo spinale e al tessuto muscolare di un topo. Questi artefatti cerebrali sembrano presentare attività fisiologiche simili a quelle che si registrano nel cervello dei neonati prematuri del settimo o ottavo mese di gravidanza. Lo studio degli algoritmi capaci di descrivere gli stati di coscienza di questi organoidi cerebrali è già al centro dell'interesse di Microsoft per lo sviluppo di sistemi artificiali che funzionino come la coscienza umana.

Termini come *machine learning* o *deep learning* sono ormai entrati nel nostro linguaggio corrente e nelle nostre applicazioni tecnologiche per indicare un fenomeno in cui macchine sempre più complesse sono in grado apprendere e rielaborare un numero sempre più ampio di dati e, nel *cloud*, un fenomeno si lega all'altro con un effetto moltiplicatore della potenza dei sistemi di calcolo. «Siamo entrati nell'era degli algoritmi, e il mondo si è ritrovato ad adorare degli indovini digitali»<sup>34</sup>.

#### 4. UMANO, NON TROPPO UMANO

In corrispondenza con l'ambizione tecnologica di emulare il cervello umano per elaborare algoritmi pensanti, è già da tempo in atto un singolare processo

31. Taddei Elmi, 2020, p. XV.

32. Bodei, 2019, p. 299.

33. Reardon, 2020.

34. Vespignani, Rijntano, 2019, p. 67.

di umanizzazione di cui è espressione tanto l'ipotesi di configurare una sorta di personalità e dignità «numerica» in analogia con la personalità e dignità umana, quanto l'aspirazione a elaborare algoritmi morali che dovrebbero guidare i processi di automazione. In una sorta di sentimentalismo tecnologico gli algoritmi sono stati accusati di essere razzisti oppure sessisti. In alcuni casi, come abbiamo visto, appaiono addirittura «stressati». McEwan sembra intercettare queste visioni quando pone al suo romanzo sulla robotica un titolo, *Machine Like Me and People Like You*, che non lascia dubbi sul fatto che vediamo le macchine sempre più come una proiezione della nostra identità. Emerge, quindi, l'esigenza di sviluppare un «umanesimo digitale»<sup>35</sup> o un «ambientalismo etico digitale»<sup>36</sup> nel quadro di un algor-etica che riesca a conciliare le nuove tecnologie con i valori fondamentali della natura umana.

Esigenze condivisibili ed auspicabili che dovrebbero, però, tener conto della situazione paradossale che si sta determinando: elaboriamo algoritmi sempre più sofisticati per sopperire ai limiti della natura umana fino all'ipotesi estrema di un algoritmo «definitivo»: «se esiste, è in grado di dedurre dai dati tutto il sapere di questo mondo: passato, presente e futuro»<sup>37</sup>. Nello stesso tempo vorremmo caricarli di umanità. Come aveva sottolineato Turing, l'algoritmo infallibile è quanto di più lontano da un essere umano possa esistere. La perfezione esclude l'umanizzazione e l'umanizzazione esclude la perfezione.

Mentre rincorriamo la speranza, o cediamo all'illusione, di plasmare eticamente gli algoritmi, di individuare codici etici da tradurre in codici macchina, non ci rendiamo conto che, forse, sta avvenendo proprio il contrario, siamo noi ad essere plasmati dai processi di codificazione algoritmica. Se «le soluzioni tecnologiche inducono e pianificano le forme semantiche»<sup>38</sup>, non possiamo escludere che l'iniziale processo di assimilazione delle nostre condotte stia diventando con il tempo un sottile processo di omologazione attraverso la sistematica circolarità di economia digitale, *social media*, informazioni programmate *ad hoc*, *clusterTech*. Le varie tecnologie *smart* inseriscono tra gli esseri umani e se stessi, oltre che tra gli esseri umani e le cose, una rete di minuscoli sensori e attuatori che seguono, registrano, renderizzano dati fisici, singoli comportamenti, specifiche condizioni ambientali. La «misurazione» del sé avviene attraverso *smart phone*, *tablet* oppure le svariate «applicazioni» annesse a dispositivi dotati di sensori. Possiamo registrare e monitorare, in modo continuo prestazioni fisiche, aspetti specifici del funzionamento del corpo (ad es. frequenza cardiaca, pressione arteriosa, metabolismo, livelli ormonali), condizioni di stress, abitudini alimentari.

35. Nida-Rümelin, Weidenfeld (2018), trad. it. 2019.

36. Floridi, 2020, p. 142.

37. Domingos, 2016, p. 16.

38. Mezza, 2018, p. 74.

Anche la casa *smart* è il nuovo orizzonte in cui si incontrano confort e controllo sociale. Saranno intelligenti il frigorifero, il cestino, i vestiti, le banconote, gli elettrodomestici, le pareti, i tappeti, gli orologi, le automobili... registrando e conservando tutte le informazioni sulle nostre azioni. Il rapporto con le cose diventa lo strumento con cui siamo analizzati, catalogati, omologati. Quali di queste informazioni sono rilevanti? E per chi? Con quanti oggetti interagiamo? Con quanti sistemi operativi siamo, direttamente o indirettamente, connessi? Quanti di questi oggetti e sistemi interagiscono tra loro, oltre che con noi? Si parla di internet delle cose per indicare quella pluralità di combinazioni sistematiche tra sensori ed algoritmi che non si limitano a registrare e analizzare i nostri comportamenti, ma li guidano e influenzano. Una fitta rete di piccole o grandi «intelligenze» che rispondono ai nostri stimoli, assimilano i nostri impulsi, interpretano e indirizzano i nostri bisogni.

Il corpo, il nome, il rango sociale, l'immagine, la firma, gli impulsi sono innegabilmente aspetti dell'identità del soggetto, ma sono altrettanto innegabilmente oggetti, che spendiamo nella vita di relazione e che quindi non ci possono appartenere in maniera esclusiva. La tecnica non rende soltanto evidente questa situazione, ma moltiplica gli strumenti e le occasioni di identificazione, sottraendoli alla sfera dei rapporti personali, che in qualche modo dipendono ancora dalle scelte individuali, per affidarli alla neutralità e indifferenza di un *data base* che segue logiche proprie. L'algoritmo è il simbolo di questo meccanismo di indifferenziazione che si sviluppa da sé, appiattendolo l'identità nella ricorsività del calcolo e intanto riconducendo ossessivamente a essa tutte le tracce che lascia, quasi condannandola a essere nulla più delle sue tracce: il codice a barre del polpastrello, del padiglione auricolare, della voce, dei tessuti biologici...

Ne è derivato uno specifico movimento, il *Quantified Self Movement*, i cui membri si sforzano di migliorare la comprensione di sé attraverso tutti gli strumenti offerti dalla tecnologia, ma pongono anche il problema inverso se non sia la controfigura digitale a plasmare il nostro modo di vivere, costruendo un mondo di sentimenti sintetici in cui sono gli algoritmi a influenzare le nostre convinzioni e a determinare le nostre emozioni. Dai robot sociali a quelli per persone anziane o disabili e fino ai sex robot: «le macchine vogliono che continuiate a usarle. In questo modo possono estorcervi altri dati così da vendere sempre di più a voi e agli altri. Le macchine ci trattengono in un mondo di macchine»<sup>39</sup>.

Come simbolicamente espresso dal titolo di un famoso romanzo di Eggers, *The Circle*, una società della trasparenza è una società dell'invadenza: «ecco la massima trasparenza. Senza filtri. Vedere tutto. Sempre»<sup>40</sup>. Senza tracce digi-

39. Turkle, 2019, p. 26.

40. Eggers, 2014, p. 56.

tali non c'è identità, non si aprono spazi sociali, non sono possibili relazioni e intanto queste tracce diventano oggetto di mercato nelle banche dati e strumento di profilazione nei *big data*. Tra privato e pubblico si inserisce l'indecifrabile orizzonte del *social* che è una socialità tecnica, artefatta, costruita da agenti privati con fini privati non sempre noti, spesso non dichiarati o addirittura non dichiarabili, in cui ciascuno di noi è immerso, di cui ha bisogno e da cui vorrebbe essere parte integrante. Qualunque dato passi per internet, e non credo vi sia oggi più alcun dato che non passi per internet, può essere intercettato (e probabilmente viene intercettato) catalogato e conservato. Fino a che punto tutto questo non determina una forma di omologazione che condiziona le nostre scelte, orienta i nostri gusti e decide persino molti dei nostri rapporti sociali? Ancora Turkle osserva che «è proprio questa la sfida attuale: combattere l'idea che sia normale perdere la propria privacy per un senso di socialità. Nel mondo di *Insieme ma soli*, consideravo i robot sociali come un terreno di prova, un luogo in cui avremmo permesso delle relazioni ingiustificate con l'inanimato, dove avremmo permesso delle relazioni che diminuiscono la nostra umanità»<sup>41</sup>.

In realtà, non sappiamo fino a che punto siano gli algoritmi ad essere influenzati dagli utenti o gli utenti a essere condizionati dagli algoritmi. Il problema è che potremmo porre questa domanda solo a un algoritmo e sarebbe sempre un algoritmo a fornire la risposta. Anche per evitare il rischio di un default tecnologico si potrebbero attivare controlli automatici per individuare attività sospette. Cambiamenti improvvisi delle rappresentazioni o dei valori potrebbero far scattare processi di revisione. Tuttavia è sempre a un algoritmo che dovremmo affidarci per controllare le alterazioni del sistema o individuare configurazioni inaspettate. In ogni caso la nostra dipendenza dagli algoritmi finirebbe per crescere in un processo circolare che dagli algoritmi ci rimanda ad altri algoritmi.

Il limite del tecno-umanesimo, se possiamo chiamare così il tentativo di umanizzare gli algoritmi, è costituito dall'autoreferenzialità che rende opache, se non imperscrutabili, le premesse cognitive. È giusto affermare che «non si sta considerando se si possa decidere *solo* come una macchina, bensì se decidere secondo un procedimento assimilabile a quello di una macchina sia in linea di principio accettabile...»<sup>42</sup>. Non dobbiamo temere l'intelligenza artificiale in quanto tale, ma l'autoreferenzialità nei limiti in cui rende impossibile confutare il profilo digitale, sfuggire ai condizionamenti delle controfigure digitali, preservare la sfera privata dinanzi alla sistematica quantificazione del sé, evitare i sortilegi al silicio degli algoritmi predittivi, sottrarsi ai condizionamenti delle emozioni sintetiche.

41. Turkle, 2019, p. 23.

42. Punzi, 2019, p. 323.



## 5. LA GIUSTIZIA DEGLI ALGORITMI

L'esperienza giuridica si è trovata di fronte al problema dell'autoreferenzialità, appena ha iniziato a impiegare i sistemi digitali di calcolo nelle decisioni giudiziali<sup>43</sup>. Un'ormai notissima sentenza della Corte Suprema del Wisconsin<sup>44</sup> doveva valutare la legittimità costituzionale dei c.d. *algoritmi predittivi* o *polizia predittiva*, impiegati già in diversi Stati americani per individuare il *risk assessment*, il rischio di reiterazione di un reato e, quindi, più in generale, la pericolosità sociale di un soggetto. La Corte non ha solo ritenuto che non costituisse una violazione dei diritti fondamentali affidarsi a un algoritmo per la valutazione della pericolosità sociale, ma ha statuito che non sussistesse il diritto di conoscere su quali basi fosse stata elaborata la stringa matematica che decideva sulla libertà individuale.

Sappiamo che questo algoritmo si chiama *Compas* e genera dei punteggi per valutare la probabilità di reiterazione di un reato sulla base di un questionario redatto dallo stesso imputato. È chiaro che la decisione avviene in base a elementi statistici, ma l'intelligenza artificiale «non dice cosa non dice», non spiega in che modo i fatti diventano numeri e i pregiudizi giudizi. Come vengono selezionati i dati? Come sono collegati? In base a quali parametri si traducono in una stringa numerica? I giudici non hanno affrontato questi problemi, si sono limitati, nel bilanciamento tra la tutela commerciale del segreto industriale e la tutela dei diritti di una persona a essere giudicato attraverso un procedimento trasparente, a privilegiare il mercato. È stata applicata la normativa sul segreto industriale (*Trade secrets*) che vieta per fini commerciali di divulgare le logiche che stanno alla base dei singoli algoritmi e, di riflesso, impone il dovere di accettare incondizionatamente gli effetti della loro applicazione. In effetti, nell'attuale economia digitale «come questi algoritmi vengono mischiati per ottenere risultati migliori è il segreto di ogni azienda»<sup>45</sup>.

I giudici non hanno fatto altro che riaffermare l'imperscrutabilità del momento fondativo di ogni forma di autorità, che è tale perché è effettiva e non perché è giusta. L'unica differenza è che le ragioni economiche prendono il posto della ragion di Stato, con il *Trade secret* che sostituisce gli *arcana imperii*. L'imperscrutabilità determina l'intangibilità del circuito autoreferenziale. Oltre al problema della trasparenza dello strumento algoritmico in sé, sussiste anche il problema della trasparenza del più ampio processo decisionale organizzativo che incorpora l'uso di un algoritmo. Molto dipende anche dal contesto in cui l'algoritmo viene utilizzato e dal tipo di decisione che viene assunta. Al di là della tutela del segreto industriale, dobbiamo pren-

43. Sulle prospettive giuridiche: Viola, 2019; Taddei Elmi, Contaldo, 2020; Alpa, 2020.

44. *State v. Loomis* 881 N.W.2d 749 (Wis. 2016).

45. Vespignani, Rijntano, 2019, p. 74.

dere atto del fatto che, dati gli attuali limiti tecnologici, è irrealistico aspettarsi che i sistemi di supporto decisionale di apprendimento automatico siano in grado di generare, in tutte le circostanze, spiegazioni complete per le previsioni che fanno<sup>46</sup>.

Forse il limite non è soltanto tecnico, ma anche matematico. Per essere imparziale, e quindi giusto, il sistema predittivo deve garantire la parità predittiva (il corretto raggruppamento dei reati commessi da un certo tipo di popolazione; in questo caso, dai bianchi e dai neri, ma già questa scelta è un pregiudizio) e l'uguaglianza di falsi positivi (il numero dei condannati ingiustamente in un gruppo deve essere tendenzialmente uguale a quello dell'altro). Ma questa parità predittiva è matematicamente impossibile perché il numero complessivo di reati commessi dalle singole categorie di persone incide indifferentemente tanto sugli innocenti quanto sui colpevoli, gravando i primi delle colpe dei secondi. «Non è possibile soddisfare i due tipi di imparzialità contemporaneamente. Anzi, in modo più generale è possibile dimostrare che esistono altre due dozzine di plausibili definizioni di imparzialità, che purtroppo in molti casi sono mutualmente esclusive»<sup>47</sup>. Se i neri commettono più reati dei bianchi, ogni nero avrà maggior probabilità di essere stigmatizzato ingiustamente. Finché ciascun individuo resterà un'ipotesi statistica dove solo per il fatto del colore della pelle rientra in uno schema, integra un processo numerico, ogni nero sarà sempre più colpevole di un bianco. Sarà condannato dalla statistica prima che dalle sue azioni.

Sulla stessa linea è anche un'altra analisi che individua il nodo cruciale nel fatto che con l'intelligenza artificiale si passa da procedure che descrivono matematicamente dei modelli, come nella normale programmazione, alla modellizzazione di fenomeni cui si collegano e aggregano milioni di variabili che altrimenti sarebbero troppo diverse per essere assimilate in uno schema unico. In questo modo la probabilità statistica altera il mero calcolo deterministico. «Proprio grazie al fatto che si tratta di modelli statistici si possono modellizzare fenomeni non deterministici che altrimenti potrebbero essere descritti solo assai grossolanamente. E possono essere trattate enormi quantità di dati che altrimenti potrebbero essere esaminate solo in coorti più limitate»<sup>48</sup>. Tuttavia più sono numerosi ed eterogenei i fattori aggregati e più cresce la disparità predittiva.

Quintarelli fa proprio l'esempio degli algoritmi predittivi adottati dalle forze dell'ordine o impiegati in alcuni procedimenti giudiziari: «i falsi positivi sono i casi di errore in cui un innocente viene erroneamente classificato come presunto colpevole. I falsi negativi sono i casi di errore in cui un colpevole

46. Babuta, Oswald, Rinik, 2018, pp. 17-18.

47. Vespignani, Rijtano, 2019, p. 94.

48. Quintarelli, 2019, p. 139.

viene erroneamente ritenuto innocente. La coperta è corta: ridurre i falsi negativi ci porta pressoché inevitabilmente ad aumentare i falsi positivi»<sup>49</sup>. L'etica dell'intelligenza artificiale riposa sulla giustizia dei rilievi statistici, ma i rilievi statistici hanno sempre un margine di astrazione e di arbitrio che li allontana dalla giustizia.

L'esperienza maturata in questi anni evidenzia, in diversi contesti, questo limite strutturale. L'algoritmo, utilizzato nel Regno Unito per stabilire i voti agli esami di maturità nel 2020, è stato ritirato dal Ministero dell'istruzione perché giudicato discriminatorio. Nel settore privato tra il 2016 e il 2020 la percentuale di aziende che ricorrono a sistemi di intelligenza artificiale per scremare le candidature per l'assunzione in un posto di lavoro sono passate dal 10% al 39%, ma non si è riusciti a eliminare il sospetto del carattere discriminatorio degli algoritmi adottati. Si propone da tempo di individuare un sistema di *auditing*, un processo di verifica sistematica dei parametri di sviluppo della programmazione. Tuttavia questa pratica dell'*auditing* riflette tutti i limiti di cui ho parlato: pur avendo accesso a tutte le componenti dell'algoritmo (il codice e i dati che utilizza) è estremamente difficile capire come si arriva a determinati risultati.

Il problema è, in qualche modo, aggravato dal fatto che agli algoritmi si chiede anche di valutare, nella selezione del personale, oltre alle capacità anche il senso morale. Si utilizzano a questo scopo particolari videogiochi (*AI-based hiring tools*): ad esempio, l'abilità nel gonfiare un determinato numero di palloncini. Ogni clic su «Pompa» fa guadagnare cinque centesimi virtuali, ma se il palloncino scoppia prima che si possa premere «Raccogli», si perde tutto. Videogiochi di questo genere, forniti da una società denominata Pymetrics, sono utilizzati da McDonald's, Boston Consulting Group, Kraft Heinz, o Colgate-Palmolive, per valutare generosità, prudenza, ambizione, attenzione. Sentimenti al silicio? Chi può escludere che tali sistemi non siano utilizzati anche per profilare la nostra personalità attraverso i tanti giochini che ci vengono offerti continuamente *on line*? Crediamo di giocare da soli e invece siamo sotto esame, stiamo mettendo in discussione una parte della nostra credibilità e ci giochiamo una parte della nostra identità?

Va, infatti, tenuto presente che tutti questi sistemi di selezione predittiva possono implicare un'invasione a vari livelli della sfera intima. Ad esempio il Tribunale Distrettuale dell'Aia ha deciso che Syri69, un algoritmo sviluppato dal Ministero degli Affari sociali olandese nel 2014 per scoprire chi avesse maggiori probabilità di commettere frodi commerciali, fosse gravemente lesivo del diritto alla *privacy* e ne ha imposto il ritiro.

49. *Ibid.* Sui limiti dei sistemi (*linear, nonlinear, polynomial, radial basis function, sigmoid, kernel*) di elaborazione che consentono di catturare dati sempre più complessi per distinguere i livelli pericolosità sociale cfr. Dressel, Farid, 2018.

Il problema etico si incrocia, quindi, con un problema tecnico. È necessario elaborare programmi di intelligenza artificiale che siano interpretabili e comprensibili (*explainable AI*) e quindi che trovino il modo di sfuggire all'opacità dell'autoreferenzialità<sup>50</sup>. Più in generale si parla di un diritto alla realtà, suggerendo di rafforzare il dovere di trasparenza con il principio di giustificabilità (*explicability*) che impone di dimostrare di avere assunto tutte le cautele necessarie per garantire la tutela dei diritti fondamentali, sul modello della *privacy by design* dell'art. 25 del Regolamento UE 2016/679.

Il diritto alla realtà si collega al principio del «controllo umano significativo» (*principle of Meaningful Human Control* – MHC) inizialmente suggerito dall'European Group on Ethics in Science and New Technologies<sup>51</sup> per l'impiego militare, ma che dovrebbe essere vincolante per tutti i sistemi informatici. Questo principio non ha soltanto un valore ideale, ma implica un radicale cambiamento di prospettiva, ridimensionando proprio il ruolo del mercato in alcune scelte di fondo. Lo ridimensiona perché pone un limite significativo all'autonomia progettuale delle imprese private, presupponendo indirettamente un efficace sistema di controlli. Lo ridimensiona perché incide proprio sul segreto industriale o commerciale. Come evidenziato dal caso *Loomis*, se i programmatori e le imprese che li finanziano possono chiudersi entro la sfera del segreto, è difficile ipotizzare che sia possibile verificare quanto margine di controllo umano sussista nelle singole macchine. Per non parlare dell'impiego militare in cui la segretezza è giustificata anche da motivi di sicurezza nazionale. L'ordine pubblico e l'ordine economico ormai si saldano in un inestricabile groviglio di istanze pubbliche e partecipazioni private in cui è difficile capire dove finisca lo Stato e dove inizi il mercato o, forse meglio, dove il mercato non si nasconda dietro lo Stato<sup>52</sup>.

Un leggero segnale di svolta è contenuto nell'art. 22 del Regolamento europeo sulla protezione dei dati 2016/679 che non mette esplicitamente in discussione la rigidità delle norme sul segreto industriale, ma dispone che «l'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona». Non è ancora l'affermazione generale di un controllo umano significativo su ogni procedimento automatizzato, ma la prima, timida, manifestazione di un diritto individuale a una decisione che non sia integralmente rimessa alle macchine. Il Considerando 71 dello stesso Regolamento parla di un vero e proprio «diritto all'intervento umano e a ottenere una

50. Quintarelli, 2020, p. 97.

51. European Group on Ethics in Science and New Technologies, *Artificial Intelligence, Robotics and 'Autonomous' Systems*, 2018.

52. Cfr. Rampini (2014, cap. I, parte II).

spiegazione della decisione conseguita», quando ci si trova di fronte a misure che valutano aspetti personali o incidono significativamente sulla persona.

Bostrom, nell'elaborare gli scenari che ci attendono dinanzi a una Superintelligenza guidata dalle nuove tecnologie, ritiene sia indispensabile elaborare un'analisi strategica per far fronte a tutta un serie di «considerazioni cruciali»<sup>53</sup>. La domanda cruciale che emerge dall'esperienza di questi anni è fino a che punto la tutela dei diritti fondamentali, l'unica strategia che il diritto conosca dinanzi all'arbitrarietà di qualsiasi forma di potere, abbia ancora senso nel *black box world*, nel mondo dell'*infinite computing*, del *quantified self*, degli «algoritmi astrattivi». Un mondo in cui ci affidiamo a calcoli sempre più complessi senza sapere dove ci possono condurre.

#### RIFERIMENTI BIBLIOGRAFICI

- Alpa, G., a cura di (2020). *Diritto e intelligenza artificiale*. Pacini.
- Babuta, A., Oswald, M., Rinik, C. (2018). *Machine learning algorithms and police decision-making legal, ethical and regulatory challenges*, RUSI Whitehall Report 31-8, [https://researchportal.northumbria.ac.uk/files/21461394/Babuta\\_et\\_al\\_Machine\\_Learning\\_Algorithms\\_and\\_Police\\_Decision\\_Making\\_OA.pdf](https://researchportal.northumbria.ac.uk/files/21461394/Babuta_et_al_Machine_Learning_Algorithms_and_Police_Decision_Making_OA.pdf).
- Bass, C. (2012). We've reached infinity – so start creating. *Wired*, 22 febbraio, <https://www.wired.co.uk/magazine/archive/2012/03/ideasbank/weve-reached-infinity>.
- Bodei, R. (2019). *Dominio e sottomissione. Schiavi, animali, macchine, Intelligenza Artificiale*. Il Mulino.
- Bostrom, N. (2018). *Superintelligenza. Tendenze, pericoli, strategie*. Bollati Boringhieri.
- Byung-Chul, H. (2014). *La società della trasparenza*. Nottetempo.
- Cachola, I., Lo, K., Cohan, A., Weld, D.S. (2020). *Findings of the Association for Computational Linguistics: EMNLP 2020*, November 16-20.
- Data Ethics Commission, German Federal Ministry of Justice and Consumer Protection (2019). *Opinion on data and algorithmic systems, including Artificial Intelligence*.
- Domingos, P. (2016). *L'algoritmo definitivo. La macchina che impara da sola e il futuro del nostro mondo*. Bollati Boringhieri.
- Dressel, J., Farid, H., (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4.
- Dumézil, G. (1935-36). Temps et mythes. In *Recherches philosophiques*, V (pp. 235-251). Vrin.
- Durante, M. (2019). *Potere computazionale L'impatto delle ICT su diritto, società, sapere*. Meltemi.
- Eggers, D. (2014). *Il cerchio*. Mondadori.
- European Group on Ethics in Science and New Technologies (2018). *Artificial Intelligence, robotics and 'autonomous' systems*, [https://ec.europa.eu/info/news/ethics-artificial-intelligence-statement-ege-released-2018-apr-24\\_en](https://ec.europa.eu/info/news/ethics-artificial-intelligence-statement-ege-released-2018-apr-24_en).

53. Bostrom (2014), trad. it. 2018, p. 213.

- Floridi, L. (2017). *La quarta rivoluzione industriale. Come l'infosfera sta trasformando il mondo*. Cortina.
- Floridi, L. (2020). *Pensare l'infosfera. La filosofia come design concettuale*. Cortina.
- Foucault, M. (1967). *Le parole e le cose*. Rizzoli.
- Gentile, F. (2003). *Politica aut/et statistica. Prolegomeni di una teoria generale dell'ordinamento giuridico*. Giuffrè.
- Geymonat, G., Giorello, G. (1977), Calcolo. In *Enciclopedia Einaudi*. Einaudi.
- Hobbes, T. (1655). De corpore. In *Elementi di filosofia: L'uomo-Il corpo* (trad. it. 1972). Utet.
- Kant, I. (1784). *Naturrecht Feyerabend*. In *Lezioni sul diritto naturale* (trad. it. 2003). Bompiani.
- Kenyon, G. (2020). Lack of Sleep Could Be a Problem for AIs. Some types of artificial intelligence could start to hallucinate if they don't get enough rest, just as humans do. *Scientific American*, 5 Dec., <https://www.scientificamerican.com/article/lack-of-sleep-could-be-a-problem-for-ais/>.
- Kurzweil, R. (2008), *La singolarità è vicina*. Apogeo.
- La Mettrie, J.O. de (1747). *L'homme machine*. In *Opere filosofiche* (trad. it. 1978). Laterza.
- Luhmann, N. (2013). *Esistono ancora norme indispensabili?* Armando.
- Marsilio da Padova (1324). *Defensor pacis*. In *Il difensore della pace* (trad. it. 1960). Utet.
- Mezza, M. (2018). *Algoritmi della libertà. La potenza del calcolo tra dominio e conflitto*. Donzelli.
- Nida-Rümelin, J., Weidenfeld, N. (2019). *Umanesimo digitale. Un'etica per l'epoca dell'intelligenza artificiale*. Franco Angeli.
- Peres, E. (2020). *Che cosa sono gli algoritmi*. Salani.
- Punzi, A. (2019). Judge in the Machine. E se fossero le macchine a restituirci l'umanità del giudicare? In A. Carleo (a cura di), *Decisione robotica* (pp. 305-316). Il Mulino.
- Quintarelli, S. (2019). *Capitalismo immateriale. Le tecnologie digitali e il nuovo conflitto sociale*. Bollati Boringhieri.
- Quintarelli, S., a cura di (2020). *Intelligenza artificiale. Cos'è davvero, come funziona, che effetto ha*. Bollati Boringhieri.
- Rampini, F. (2014). *Rete padrona. Amazon, Apple, Google & co. Il volto oscuro della rivoluzione digitale*. Feltrinelli.
- Reardon, S. (2020). Can lab-grown brains become conscious? *Nature*, 4 Nov., <https://www.nature.com/articles/d41586-020-02986-y>.
- Schmitt, C. (1936-7). *Scritti su Thomas Hobbes* (trad. it. 1986). Giuffrè.
- Seung, S. (2013). *Connettoma. La nuova geografia della mente*. Codice.
- Taddei Elmi, G. (2020). Introduzione. Dall'informatica giuridica al diritto dell'informatica. In Id., *Algoritmi giuridici. Ius condendum o "fantadiritto"*. Pacini.
- Turkle, S. (2019). *Insieme ma soli. Perché ci aspettiamo sempre più dalla tecnologia e sempre meno dagli altri*. Einaudi.
- Valerio, C. (2020). *La matematica è politica*. Einaudi.
- Vespignani, A., Rijtano, R. (2019). *L'algoritmo e l'oracolo. Come la scienza predica il futuro e ci aiuta a cambiare*. Il Saggiatore.

#### EMOZIONI SINTETICHE E SORTILEGI AL SILICIO

- Vineis, P., Carra, L., Cingolani, R. (2020). *Prevenire. Manifesto per una tecnopolitica*. Einaudi.
- Viola, L., a cura di (2019). *Giustizia predittiva e interpretazione della legge con modelli matematici. Atti del Convegno tenuto presso l'Enciclopedia Italiana Treccani*. Centro Studi Diritto Avanzato.
- Zuboff, S. (2019). *Il Capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*. Luiss University Press.

