INTRODUZIONE

Nel 1997, il programma Deep Blue, sviluppato dai ricercatori della Carnagie Mellon University, sconfisse il campione mondiale di scacchi Garry Kasparov. Il campione non accettò la sconfitta e, sul presupposto che una macchina non avrebbe mai potuto elaborare strategie così intelligenti da superare l'intelligenza umana, accusò gli scienziati di aver truccato la gara. Questa vittoria mise in discussione il primato della mente umana nei confronti delle macchine, infatti gli scacchi erano stati a lungo la fortezza dell'intelligenza umana che l'automazione non aveva ancora mai espugnato[1]. Oggi, i programmi capaci di giocare a scacchi sono talmente potenti da non venire più testati contro avversari umani ma contro computer. Questo è ciò che chiamiamo Intelligenza Artificiale (AI), un'intelligenza che nacque negli anni Cinquanta del XX secolo e che è parte integrante della rivoluzione digitale, la quale ha comportato e comporterà numerose trasformazioni sociali.

Ma una forma embrionale di AI era apparsa già nei primi anni '40, quando Alan Turing, il padre dell'informatica, violò i codici di Enigma, la macchina che i tedeschi usavano per inviare messaggi segreti indecifrabili per gli Alleati.

Oggi, i progressi fatti dall'AI, accompagnati dalla robotica, hanno permesso la nascita di sistemi capaci di competere con le capacità umane e, in alcuni casi, di superarle. Questi sistemi sono in buona parte autonomi poiché capaci di imparare dalle loro stesse esperienze, attraverso il *machine learning* e di migliorare grazie ai dati immagazzinati. Sono, per queste ragioni, capaci di compiere azioni che i loro stessi ideatori non avevano programmato, tanto che, oggi, non ha più senso dire che i computer fanno solo quello che sono programmati per fare.

Il progredire delle capacità intellettuali dell'AI sta cambiando il modo di vivere degli esseri umani e rivoluzionerà il mondo del lavoro. Vi sono sistemi capaci di pensieri e

---

[1] Altra tappa nel superamento dell'uomo da parte delle macchine fu quella segnata dalla vittoria al quiz televisivo *Jeopardy* da parte del programma IBM Watson, nel 2011, che utilizzò un database di 200 milioni di pagine di fatti e cifre, compresa l'intera Wikipedia, per un totale di quattro terabyte di memoria. Successivamente, nel 2015, una nuova tappa nella storia dell'AI fu segnata dalla vittoria di AlphaGo, un'AI sviluppata da Deep-Mind, di proprietà di Google, sul campione Fan Hui, uno dei tre migliori giocatori di Go al mondo. La differenza tra il gioco degli scacchi e il gioco del Go è che il primo obbedisce ad una logica matematica, pertanto razionale, mentre l'altro si fonda su una logica intuitiva. È più semplice inserire nella memoria di un computer tutte le combinazioni possibili del gioco degli scacchi, laddove è molto più difficile creare artificialmente l'universo mentale del gioco del Go. Per queste ragioni la vittoria di AlphaGo, che all'epoca gli esperti non si aspettavano prima di 10 o 20 anni, ha segnato una svolta nella storia dell'AI.

azioni autonome che pongono il problema di individuare il soggetto responsabile per tali azioni. I sistemi intelligenti pongono, dunque, notevoli quesiti, come quello di capire se è giusto che essi abbiano diritti e responsabilità individuali e fino a che punto possano spingersi senza violare diritti fondamentali dell'uomo.

I giganti dell'era dell'AI sono attualmente almeno sette: Google, Facebook, Amazon, Microsoft, Baidu, Alibaba e Tencent. Essi accostano all'AI progetti transumanisti che permetteranno all'uomo poteri illimitati, come modificare il suo genoma, riprogrammare il suo cervello e "sconfiggere" la morte. L'AI renderà davvero gli uomini simili a degli dei?

La presente disamina si propone di calibrare gli sviluppi dell'AI in relazione al suo impatto sulla società, in particolare sui diritti umani e le libertà fondamentali come il diritto alla riservatezza e alla tutela dei propri dati personali, poiché gli algoritmi, specie quelli di *machine learning* sono spesso accusati di generare iniquità, interferenze nella sfera privata, discriminazioni e opacità. Si tratta di un ambito che solo di recente ha iniziato a porre le basi per una regolamentazione legislativa in chiave etica, vedi tra tutti il nuovo progetto di Regolamento europeo per l'AI. Certi diritti sono destinati a soccombere di fronte al dilagare dell'AI o potrà l'intelligenza umana individuare dei controlimiti capaci di scongiurare i pericoli che inevitabilmente si accompagnano all'evoluzione digitale[2]?

[2] Tratto da Cataleta M.S. et al., *Artificial Intelligence and Human Rights, an Unequal Struggle*, in *CIFILE-Canadian Institute for International Law Expertise Journal of International Law*, Vol. 1, No. 2, 2020.

PREFAZIONE

Si è molto discusso negli ultimi anni di una «crisi», se non della «fine», dei diritti umani, specie sotto la pressione del c.d. «sovranismo» diffusosi a livello globale in concomitanza con le posizioni dell'*ex* Presidente degli Stati Uniti D. Trump. La questione va ricondotta, più indietro nel tempo, almeno alla politica c.d. «neoliberista» della destra anglo-americana, che è riuscita a «riformulare» i «valori» occidentali, contro il blocco comunista durante gli ultimi anni della guerra fredda, in termini di efficientismo e oggettivismo economico, dando una patina di scientificità di «non-ritorno» o di «ovvia inesistenza di alternative credibili» ad una strategia di potenza e di controllo globale delle masse, ed erodendo i diritti della persona umana intesa nella sua *unicità* e *complessità* attraverso tecniche di omogeneizzazione e semplificazione consumistica applicate ad ogni aspetto della vita sociale. La c.d. «intelligenza artificiale» (IA, AI), i cui vantaggi nessuno può negare (come si può dire di *qualsiasi* progresso tecnologico), *in questo contesto* rappresenta la sintesi del *gregarismo sociale* che cerca di annullare la «persona umana» nella sua unicità e complessità, non diversamente sotto questo profilo da quanto avviene in qualsiasi regime totalitario, non importa se autocratico o democratico. Parlare di «diritti umani» non ha allora alcun senso, così come parlare della loro «crisi», se non si modifica *a livello psichico* lo sfondo autoritario che guida le tecniche pubblicitarie intese a trasformare gli stessi diritti umani in prodotti commercialmente funzionali.

Occorre registrare che si vanno moltiplicando presso organizzazioni internazionali e altre istituzioni (ad esempio nuovi corsi di laurea nelle università) gli «studi» sulla IA. La locuzione è quanto mai rivelativa del suo intento di fondo, che non è certo l'intelligenza, comunque la si intenda, ma è l'artefatto che dà maggiore *potenza* a chi se ne serve, al punto che l'epoca attuale potrebbe descriversi come quella della «AI Race», non dissimile dalla «Space Race» degli anni '60 del XX sec. Tra gli innumerevoli documenti internazionali elaborati più recentemente possono ricordarsi, tra quelli adottati nel 2019, le «Nuove Linee-guida sulla protezione dei dati e l'intelligenza artificiale» (Consiglio d'Europa), la risoluzione sulla «politica industriale europea globale in materia di robotica e intelligenza artificiale» (Parlamento UE), una raccomandazione sull'intelligenza artificiale diretta «a stimolare l'innovazione» (OCSE); e, tra quelli adottati nel 2020, la «Carta bianca sull'intelligenza artificiale» (Commissione UE), la raccomandazione sull'«etica dell'intelligenza artificiale» (UNESCO) e una raccomandazione relativa all'«impatto dei sistemi algoritmici sui diritti umani» (Consiglio d'Europa); da ultimo, il 21 aprile 2021, la Commissione UE ha

pubblicato una proposta di regolamento sull'intelligenza artificiale. In tutti questi documenti l'*incipit* di *routine* è immancabilmente che l'IA è molto utile ma può comportare dei rischi. L'accento viene poi dosato diversamente a seconda del contesto e di solito le organizzazioni economiche esaltano i benefici, pur notando di passaggio i rischi, mentre le organizzazioni umanitarie appaiono più attente ai possibili effetti negativi di tipo etico e sui diritti umani. A scorrere questi documenti, incredibilmente prolissi e ripetitivi, si nota che: (*a*) i termini «intelligenza» e «artificiale» non vengono definiti ma dati per acquisiti; (*b*) si insiste nel definire l'approccio seguito come «umano-centrico» ma senza definire in base a quale definizione di «umano»; (*c*) si raccomanda l'inculcazione della «fiducia», cioè di una fede progressivamente interiorizzata, nell'IA come catalizzatrice dei benefici promessi, che appaiono essenzialmente di tipo materiale; (*d*) si dà per scontato e *già deciso* che l'IA va perseguita, rimanendo da discutere solo il modo di muoversi così da raggiungere prima degli altri la leadership globale secondo i propri valori (ad esempio da parte dell'Unione Europea contro Stati Uniti e Cina). Si tratta di un approccio di *potenza* (e di *scontro* aperto) nel quale la posta in gioco è chi controlla il mondo, e l'IA, come da sempre la tecnologia, è cruciale. Parlare di «etica» e di diritti umani, e addirittura di un'«etica» (anch'essa *mai definita* tra le innumerevoli etiche che da millenni si contendono il campo) «dell'IA», in questo contesto può suonare retorico ai più ingenui e perfettamente funzionale al sistema di potenza ai più scaltri.

In questo suo volume Maria Stefania Cataleta descrive in modo pieno ed efficace i principali problemi sollevati dall'IA sul piano etico e giuridico, offrendo diversi spunti di riflessione, ad esempio quando esamina la c.d. «rivalità» tra IA e intelligenza umana e la dicotomia tra IA «debole» e «forte»; l'impatto dell'IA sia sulle società in quanto tali sia sui singoli individui e sui diritti umani (in particolare il rispetto della privacy in concomitanza con lo sfruttamento dei *big data*, la non-discriminazione e la libertà di espressione); la recente azione europea (sia dell'UE sia del Consiglio d'Europa) a favore di un equilibrio tra benefici e rischi dell'IA; la difesa della democrazia da usi perversi dell'IA; la geopolitica dell'IA quale strumento chiave di potenza globale nella sfida (soprattutto, ma non solo) tra UE, Stati Uniti e Cina. Quest'ultimo punto è messo in evidenza, ad esempio, quando l'autrice sottolinea che «stiamo diventando una colonia digitale dell'industria dell'AI americana e cinese» e che «Stati Uniti e Cina sono consapevoli che l'AI rappresenta un enorme vantaggio competitivo». Né mancano gli (importanti) aspetti psicologici, come quando l'autrice osserva che «I sistemi di *machine learning* sono attualmente in grado di influenzare emozioni subliminalmente». Alla luce di così numerose variabili, l'autrice conclude che «È impossibile applicare un criterio precauzionale troppo rigido, poiché questo inficerebbe la possibilità per il singolo di

beneficiare del progresso tecnologico, laddove è più ragionevole un bilanciamento tra l'indubbio vantaggio derivante dall'applicazione intelligente e un margine di rischio tollerabile per l'utente».

Da europei non possiamo che augurarci che l'UE riesca nel suo compito di assicurare diritti umani demercificati e deautomatizzati e contenere quel gregarismo di potenza che passo passo può farci riprecipitare nell'abisso della guerra.

Carlo Focarelli
Professore Ordinario di Diritto Internazionale
Università Roma Tre