



Elena Falletti

# Discriminazione algoritmica

Una prospettiva comparata

*Prefazione di* Roberto Pardolesi



**Giappichelli**

## PREFAZIONE

Il libro di E.F. è, innanzi tutto, una straordinaria miniera di informazioni. Troppe, ecciperà qualcuno. E potrebbe anche essere vero, per chi parta dal postulato che ogni sforzo autenticamente scientifico deve mirare alla riduzione della complessità. Non è meno evidente, però, che la ‘quarta rivoluzione’, per usare la suggestiva tassonomia di Luciano Floridi, non si lascia catturare da una ricostruzione monotonica. Laddove i sommovimenti precedenti avevano registrato il declino dell’uomo come figura centrale nel progetto divino e nello svolgersi dei fenomeni cosmologici a seguito della rivoluzione copernicana, il passaggio ad un’idea dell’essere umano come esito di un processo evolutivo analogo a quello delle altre forme di vita, determinato principalmente dall’interazione gene/ambiente propria della riflessione condotta da Darwin, sino alla rimozione dell’io, imposta dal pensiero freudiano, dal “centro di se stesso”, la rivoluzione digitale rappresenta la più radicale (e potenzialmente devastante) alterazione di paradigmi, perché rimuove la centralità dell’uomo financo rispetto al proprio processo decisionale: non più esito della ragione quanto di forze dinamiche d’impulso, da leggersi in ottica meccanicistica, spogliandolo della prerogativa di unico decisore “logico” a seguito degli sviluppi della tecnologia e del manifestarsi di complessi automatici che, coerentemente con quanto il test di Turing richiedeva, sono in grado di simulare mimeticamente il comportamento umano, rendendo impossibile operare una qualche significativa distinzione. Orbene, questa rivoluzione non si lascia ridurre alla sola dimensione del potere computazionale, capace di gestire a costi viepiù contenuti masse sconfinite e sempre crescenti di dati, sino a propiziare l’equivoco dello scambio di quantità per qualità nella riproduzione dei processi decisionali, nel presupposto (illusorio) che la sommatoria dei dati possa fornire qualcosa di più di una traccia inferenziale, meramente stocastica. C’è dell’altro, destinato a sfociare nella consapevolezza che l’elaborazione soltanto sintattica del dato diaforico si presta a gestire la routine (il più delle volte saranno rieditate condotte pregresse), ma non è comunque in grado di soppiantare la sintesi semantica dell’attribuzione di valore. E ciò contribuisce ad arricchire un quadro che non può contentarsi di scelte manichee, ma richiede un esame olistico, che di quella complessità/compenetrazione di aspetti sappia fare tesoro. Ecco, allora, che la sapiente ricognizione comparativa che E.F. propone diventa vero banco di prova per chi voglia seriamente saggiare le prospettive prossime venture dell’intelligenza artificiale e di tutto quanto vi si riconnette. Si rianno-

dano, così, molti fili; e si misurano apporti che mette conto scrutinare con estrema attenzione.

Chi è senza peccato, ammoniva una voce che non teme smentita. E infatti, quando si va al cuore della ricerca del lavoro di E.F., i pochi rilievi che ho testé riassunto si ritorcono contro di me. Perché, sul tema della discriminazione algoritmica avevo maturato – si fa per dire! – il convincimento che alla macchina, operante sulla base di un software che qualcuno ha a suo tempo programmato, anche in vista degli sviluppi da *e-learning*, non si possa chiedere più di quanto chiediamo al decisore umano. Al quale, salvi casi assai particolari, siamo soliti perdonare, soprattutto per la pratica impossibilità di impedirlo, che le sue valutazioni siano conformate dal suo bagaglio di esperienze, ossia, fuori dai denti, dei suoi pregiudizi. Sarà utile ricordare – a mo' di esempio paradigmatico – il celebre caso da cui ha tratto vigore il movimento di pensiero (giuridico) che ha portato al riconoscimento del danno biologico: il figlio di un muratore (si scrisse in una sentenza che destò addolorato scalpore), il quale abbia riportato in tenera età una grave menomazione per l'altrui fatto illecito, ha diritto al ristoro, da commisurare sul fatto che, da grande, farà, con ogni probabilità il ... muratore. Ragionamento cinico e concettualmente discriminatorio, quanto, nella sostanza, ineccepibile sul piano delle proiezioni statistiche. Si ha un bel dire sul riscatto sociale, il sogno americano e via discorrendo. Le probabilità che lo sfortunato ragazzino diventasse, nel tempo, capitano d'industria o quant'altro erano decisamente minori di quelle dischiuse al giovane virgulto di una famiglia di affermati professionisti. Non piace ammetterlo, ma sono molte le cose che non ci garbano, ivi compresa, tanto per saltare a un'altra situazione urticante, la voce ricorrente, almeno in un passato non troppo remoto, che gli uffici del personale delle grandi imprese, all'atto di scandagliare *curricula* di nuove leve a fini di reclutamento, solessero scartare sistematicamente le domande di giovani laureati provenienti da (talune) università del Sud.

Quei pregiudizi, opinavo io, sono radicati nella nostra cultura, meglio, nel nostro essere; e non c'è vero modo di sradicarli, se non mercé la constatazione che aderire a quei pregiudizi riduce, sì, i costi d'informazione, ma a prezzo di falsi positivi/negativi che comportano oneri sicuramente eccedentari. Se così è, concludo, perché mai dovremmo chiedere all'algoritmo, che abbiamo impostato noi, di andare al di là delle nostre capacità di giudizio? Perché scandalizzarsi di una discriminazione algoritmica, se l'alternativa è rappresentata dalla discriminazione di un qualche *credit bureau*?

E.F., con dovizia di argomenti, dimostra che il semplicismo del mio approccio è frutto di un'analisi superficiale. I fattori da metter nel conto sono, appunto, tanti e molto più complicati. E questo libro è un invito provvido ad approfondirli.

Roberto Pardolesi

## INTRODUZIONE

Il diritto è una disciplina che sembra assomigliare al Giano Bifronte<sup>1</sup>. Da un lato, esso si volge al passato: la norma giuridica viene applicata rispetto a un caso concreto esistito e concluso nel tempo. Dalla ricostruzione complessiva delle singole applicazioni giuridiche è possibile delineare il diritto vivente, inteso quale risultato della evoluzione storica e sociale di una comunità<sup>2</sup>. Dall'altro lato, compito del diritto è quello di disciplinare il futuro, cioè di costruire, attraverso la definizione di norme giuridiche, lo sviluppo di situazioni ancora *in itinere*.

In tale condizione di disciplina ancora incerta si trovano alcune realtà relative al rapporto tra innovazione e diritto, in particolare la relazione uomo-macchina, la quale è spinta dall'ambizione di riuscire a riprodurre artificialmente le capacità umane logiche e di ragionamento. Tuttavia, questa ambizione ha già oltrepassato il confine tra l'uomo e la proiezione meccanica dei suoi bisogni, in particolare il desiderio di efficienza e risparmio di tempo ed energie, ma al contempo è emersa con preoccupazione l'idea che un siffatto modello di automazione possa travolgere l'*humanitas*<sup>3</sup>.

Infatti, con l'espressione "intelligenza artificiale" si intende comunemente la capacità attribuita a una macchina di elaborare operazioni complesse riconducibili a quelle effettuate da un cervello umano. L'idea di costruire una macchina in grado

---

<sup>1</sup> Questa immagine aderisce bene anche all'ambito del diritto dell'informatica, poiché da un lato si rivolge verso i problemi giuridici dell'informatica, mentre dall'altro si occupa dell'uso dell'informatica nel diritto (G. Sartor, *L'informatica giuridica e le tecnologie dell'informazione*, Giappichelli, Torino, 2016, p. 14).

<sup>2</sup> Per quel che concerne la tradizione giuridica occidentale, le radici del diritto sono assai profonde e traggono linfa fin dalla tradizione classica greca e romana, per poi attingere ad esperienze come il cristianesimo e l'illuminismo (L. Holmes, *Ethics, Legal Ethics, and the Classical Tradition*, in 67 *Ark. L. Rev.* 351, 2014; R.B. Kar, *Western Legal Prehistory: Reconstructing the Hidden Origins of Western Law and Civilization*, in 2012 *U. Ill. L. Rev.* 1499; S. Todd Lowry, *The Economic and Jurisprudential Ideas of the Ancient Greeks: Our Heritage from the Hellenic Thought*, in S. Todd Lowry, B. Gordon, *Ancient and Medieval Economic Ideas and Concepts of Social Justice*, Brill, Leiden, New York, Köln, 1998, p. 11).

<sup>3</sup> N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford-New York, 2014. Vi è chi contesta fondatamente questo approccio, ammettendo però che esso ha avuto significative conseguenze culturali, in quanto ha contribuito a creare un mito che impedisce una visione neutra al tema (E.J. Larson, *The Myth of Artificial Intelligence, Why Computers Can't Think the Way We Do*, Harvard University Press, Cambridge, 2021).

di replicare le fattezze e i comportamenti umani non è affatto recente e la si può ritrovare nei racconti ereditati dal passato, come quello di Pigmaliione che costruì una statua così bella da sembrare viva e innamorarsene o del Golem praghese, ubbidiente difensore della sua comunità.

Fino a poco tempo fa l'introduzione massiva dell'automazione intelligente appariva relegata a settori non pertinenti con la quotidianità: cinema e letteratura illustravano irrealistici e, a volte, inquietanti scenari dove le macchine prendevano il sopravvento sul genere umano con conseguenze catastrofiche<sup>4</sup>. Infatti, la fantascienza ha contribuito a diffondere nell'immaginario l'ipotesi stereotipata che le macchine intelligenti saranno in grado di prevalere sugli esseri umani, dominandoli. Tale ipotesi è legata allo sviluppo di una intelligenza artificiale in grado di imitare il comportamento umano in specifici settori della vita collegati con la capacità di calcolo, come i giochi, la matematica e il linguaggio, con l'aspirazione di riprodurre, automatizzando siffatte condotte<sup>5</sup>. L'ambizioso obiettivo di tale approccio concerneva la simulazione dell'intelligenza umana affinché potesse essere riprodotta e automatizzata nelle macchine.

Questa prospettiva oggi è ormai tramontata e la ricerca in materia si concentra sulla progettazione di sistemi automatizzati in grado di elaborare risultati adeguati in domini complessi non più soltanto con l'imitazione del pensiero umano, ma con qualsiasi mezzo<sup>6</sup>, focalizzandosi sull'intelligenza come concetto astratto.

Invece, nella realtà quotidiana, l'automazione si è diffusa con maggiore discrezione: in ambito giuridico, per esempio, l'utilizzo di strumenti di automazione decisionale, quali i sistemi legali esperti, sarebbe potuto diventare strumentale allo svolgimento più efficiente del procedimento giudiziario civile<sup>7</sup>, però senza mai avere avuto un significativo utilizzo pratico.

Tempo dopo, i sistemi esperti sono stati marginalizzati sia nell'ambito della ricerca sia nell'uso professionale, mentre gli sforzi dei programmatori si sono orientati sull'applicazione del *machine learning* nelle decisioni automatizzate. Gli Automated Decision Making System (d'ora in poi – ADMs) sono diventati pervasivi in molti ambiti nella vita quotidiana. Infatti, il procedimento di *ranking* reputazio-

---

<sup>4</sup> A.D. Hudson, E. Finn, R. Wylie, *What can science fiction tell us about the future of artificial intelligence policy?*, in *AI & Society*, 2021, *passim*; A. Favaro, *Digital Culture: Fears, Hopes And Uncertainty About Artificial Intelligence In Sci-Fi Movies*, in A. Ayhan, *New Approaches in Media and Communication*, Peter Lang, Berlin, 2019, p. 61 ss. Per una ricostruzione focalizzata più sugli aspetti di comunicazione e marketing: P. Bory, *Deep new: The shifting narratives of artificial intelligence from Deep Blue to AlphaGo*, in *The International Journal of Research into New Media Technologies*, 2019, p. 1-16.

<sup>5</sup> S. Dick, *Artificial intelligence*, in *Harvard Data Science Review*, 2019, 1(1), <https://doi.org/10.1162/99608f92.92fe150c>.

<sup>6</sup> S. Dick, *Artificial intelligence*, cit.; L. Floridi, *The fourth revolution: How the info sphere is reshaping human reality*, Oxford University Press, Oxford, 2016.

<sup>7</sup> R. Susskind, *Expert Systems in Law: A Jurisprudential Approach to Artificial Intelligence and Legal Reasoning*, in *Modern Law Review*, 1986, p. 168-194.

nale, caratteristico di questi modelli, si concretizza con la valutazione del c.d. *scoring*, cioè con l'assegnazione di un punteggio, in ambito bancario, assicurativo, lavorativo, sanitario<sup>8</sup> e così via. Detto punteggio viene determinato dal procedimento di decisione automatizzata cui i soggetti interessati sono sottoposti attraverso l'elaborazione dei dati personali diffusi quotidianamente dai titolari stessi in modo inconsapevole. Accade, quindi, che i dati prodotti dagli interessati vengono utilizzati contro le loro fonti in decisioni che riguardano la vita dei medesimi.

L'evoluzione del *data management* e del *decision making* affonda le sue origini negli Anni Settanta del Secolo scorso, quando venne sviluppato un sistema decisionario definito "*Decision Support System*" (DSS), il quale era caratterizzato da un singolo elemento decisionale relativo da uno o pochi data base interni. Le capacità di tale sistema erano limitate a decisioni strategiche o tattiche, compatibilmente con la tecnologia allora disponibile<sup>9</sup>. Vent'anni dopo, negli Anni Novanta, la seconda generazione di sistemi decisionali era definita "*enterprises data warehouse*" (EDW) e vedeva l'incremento delle soluzioni ai bisogni da soddisfare, grazie all'aumento dei data base e al miglioramento dei limiti organizzativi.

L'apporto di siffatti sistemi ai loro utenti non cambiava nella natura (trattandosi di supporto per decisioni strategiche ovvero tattiche), ma era migliorata la qualità delle soluzioni proposte, nonostante la tecnologia utilizzata rimanesse ancora primitiva<sup>10</sup>. Un ulteriore miglioramento di tali sistemi si concretizzò negli Anni Zero del nuovo millennio. La terza generazione dei sistemi decisori acquisì la capacità decisionale in materia strategica divenne più strutturata e con risultati consegnati in tempo reale<sup>11</sup>.

Tuttavia, il cambiamento di paradigma in questo ambito avvenne negli Anni Dieci del nuovo secolo, con la quarta generazione dei sistemi di *decision-making*, quando la quantità di dati disponibili crebbe esponenzialmente, consentendo la creazione di "*Business Intelligence (BI) System*", i quali erano in grado di comprendere i dati, affinare le strategie sia in senso strategico, sia in senso operativo e adottare delle decisioni vere e proprie<sup>12</sup>. Con gli Anni Venti si è arrivati a un livello di gestione delle banche dati e dei sistemi decisori complesso e integrato, infatti la capacità relazione e di connessione delle fonti informative è stata raffinata attraverso l'uso dei "*Relational Database Management System*" (RDBMS), cioè tecnologie in grado di usare i *big data* al fine di creare un contesto più ricco con le informazioni provenienti da *device*, piattaforme e *cloud storage*<sup>13</sup>.

Seppur apparentemente possa sembrare che i big data abbiano travolto l'esisten-

---

<sup>8</sup> F. Chiusi, S. Fischer, N. Kayser-Bril, M. Spielkamp, *Report Automatic Society, 2020*, Algorithm Watch, Bertelsmann Stiftung, Berlin, 2020, *passim*.

<sup>9</sup> S. Gressel, D.J. Pauleen, N. Taskin, *Management Decision-Making, Big Data Analytics*, cit., p. 24.

<sup>10</sup> S. Gressel, D.J. Pauleen, N. Taskin, *op. cit.*

<sup>11</sup> S. Gressel, D.J. Pauleen, N. Taskin, *op. cit.*

<sup>12</sup> S. Gressel, D.J. Pauleen, N. Taskin, *op. cit.*

<sup>13</sup> S. Gressel, D.J. Pauleen, N. Taskin, *op. cit.*

za degli individui, in realtà si è trattato di una trasformazione avvenuta progressivamente: come una sorta di percezione capovolta, simile a quella provocata dal Gibbons quando scrisse “Declino e caduta dell’Impero Romano”<sup>14</sup>.

Anche la visione popolare e convenzionale dell’intelligenza artificiale, in particolare l’ambito delle applicazioni automatizzate in ambito decisionale, soffre di un analogo stereotipo relativo ad una idea largamente diffusa, ma che non è coerente con la realtà. Nonostante la comune convinzione che l’implementazione dell’automazione sarà repentina, essa invece riguarda la realizzazione di una lenta e silenziosa<sup>15</sup> trasformazione avvenuta con la graduale assimilazione degli individui all’uso, soprattutto passivo, dei procedimenti decisionali automatizzati.

Invece, il senso della nuova Proposta di Regolamento in materia di Intelligenza Artificiale pubblicato<sup>16</sup> dalla Commissione Europea approccia il tema come se gli strumenti che utilizzano il *machine learning* fossero in contrapposizione all’*humanitas* e il loro avvento fosse paragonabile a quella dei Cavalieri dell’Apocalisse raffigurati da Albrecht Dürer: improvvisa e devastante. Al contrario, l’ambizione umana di avvalersi di strumenti meccanici è antica e profonda, come si può dedurre dall’espressione *Deus ex machina*, dato che le macchine, gli automi di scena, erano già in uso in tempi ancestrali.

Ci si domanda, in via generale, quali siano le conseguenze dell’esposizione della personalità digitale nei confronti dell’automazione dei processi decisionali e come da tali situazioni possano emergere situazioni di discriminazione a causa della disintermediazione informativa<sup>17</sup>. Il ruolo dell’automazione dei procedimenti decisionali, giudiziari e non, è di particolare interesse in un ambito dove la regolamentazione giuridica non può essere considerata stabilmente compiuta, ma deve affrontare il continuo evolversi della tecnologia.

---

<sup>14</sup> L’autorevole storico inglese contribuì a creare uno dei più significativi, durevoli e involontari stereotipi culturali a tutt’oggi persistenti e difficili da demistificare sulle cause della rovina della civiltà romana e cioè che furono le torme barbariche a provocarne la caduta. L’Opera del Gibbons presenta una certa idea di desertificazione in conseguenza delle invasioni barbariche dopo il IV secolo d.C., mentre l’Impero romano d’Occidente non cadde a causa di improvvise e invincibili calate di popolazioni straniere, ma conobbe una lenta trasformazione che lo cambiò dall’interno (P.S. Wells, *Barbari. L’alba del nuovo mondo*, trad. it. F. Ossola, Lindau, Torino, 2008, p. 5; J.G.A. Pocock, *Edward Gibbon in History: aspects of the Text in The History of the Decline and Fall of the Roman Empire. The Tanner lectures on human values*, 1989, 11, pp. 289-364; C. Roberts, *Edward Gibbon and the Shape of History*, Oxford University Press, Oxford, 2014, *passim*).

<sup>15</sup> C.R. Sunstein, *Governing by Algorithm? No Noise and (Potentially) Less Bias*, in *Duke Law Journal*, 2022, p. 1185 ss.

<sup>16</sup> Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull’Intelligenza Artificiale (Legge sull’Intelligenza Artificiale) e modifica alcuni atti legislativi dell’Unione) pubblicato il 21 aprile 2021 e disponibile su <https://eur-lex.europa.eu/>.

<sup>17</sup> Il concetto di “disintermediazione” riguarda il processo mediante il quale vengono eliminate le strutture di mediazione, definite come “corpi intermedi” ovvero “filtri” tra due o più utenti in un procedimento complesso, il quale può riguardare una struttura di tipo pubblico, una fornitura di servizi o un processo di comunicazione (G. Riva, *Fake news*, Il Mulino, Bologna, 2018, p. 85).

In siffatta cornice, si manifesta un aspetto innovativo rispetto al passato, in quanto agli algoritmi automatizzati viene delegata<sup>18</sup>, seppure al momento in misura limitata, la funzione decisoria. Tale funzione presenta due caratteristiche tradizionalmente collegate tra loro: è propria di un giudice e la manifestazione di una forma di sovranità molto elevata, cioè il potere dello Stato di imporre sanzioni, anche estreme.

Entrambe tali caratteristiche si perdono quando l'esercizio decisorio è delegato ad un procedimento automatizzato. Tuttavia, in caso in cui il soggetto, sul quale si irradiano gli effetti di tale decisione automatizzata abbia subito lesione ingiustificata dei suoi diritti, potrà far valere il suo diritto alla spiegazione ex art. 22 GDPR e, se insoddisfatto, impugnarla di fronte agli organi giurisdizionali competenti<sup>19</sup> per criticarne i passaggi, porre in evidenza le debolezze e quindi ottenerne una soddisfacente modifica.

Diventa pertanto significativo appurare come le decisioni automatizzate vengano confutate ovvero confermate dai giudici umani sotto il profilo dell'analisi giuridica, la quale trova un riscontro motivazionale nelle sentenze dei tribunali, ma non nei risultati algoritmici.

Questo specifico aspetto riconduce la questione in parola alla nota, e probabilmente irrisolvibile, contraddizione secondo cui da un lato il giudice deve essere fedele al pensiero montesquieiano: *“les juges ne doivent être que la bouche qui prononce les paroles de la loi, des êtres inanimés qui n'en peuvent modérer ni la force ni la rigueur”*<sup>20</sup>, ma dall'altro debba altresì obbedire al dettato costituzionale secondo cui *“(T)utti i provvedimenti giurisdizionali devono essere motivati”*<sup>21</sup>.

Alla luce di ciò, e nella prospettiva di questo testo, si privilegia lo studio dei sistemi decisori automatizzati, rispetto al tema più ampio dell'intelligenza artificiale, considerando che l'intelligenza artificiale stessa è riconducibile ad una serie successiva di decisioni effettuate attraverso sistemi di calcolo matriciali e interdipendenti tra loro. In particolare, la domanda di ricerca inerente a questa trattazione è sorta dalla constatazione fattuale dell'ormai diffuso utilizzo di sistemi automatizzati raffinati e complessi che tendono a sostituire l'elemento decisorio umano, soprattutto in caso di situazioni standardizzate, ove sono presenti da un lato quantità massive di dati e dall'altro condizioni di stabilità socio-culturali, almeno apparenti, dato che sono i soggetti più vulnerabili a subire gli effetti discriminatori degli ADMs.

La difficoltà della trattazione del tema concerne la circostanza relativa all'impossibilità di focalizzare l'attenzione soltanto su un aspetto specifico. Per esempio, se ci si concentrasse esclusivamente sulla protezione dei dati personali, sfuggirebbe

---

<sup>18</sup> A. Garapon, J. Lassègue, *La giustizia digitale. Determinismo tecnologico e libertà*, trad. it. F. Morini, Il Mulino, Bologna, 2021, p. 241 ss.

<sup>19</sup> *Supra*, Cap. IV.

<sup>20</sup> C. Secondat, baron de Montesquieu, *L'Esprit des Loix*, 1748, XI, p. 6.

<sup>21</sup> Art. 111, co. 6, Cost.

l'elemento massivo del trattamento dei medesimi nel momento in cui vengono elaborati dagli algoritmi e quindi la percezione del rischio discriminatorio verrebbe eluso. Altresì, se l'attenzione venisse concentrata esclusivamente sul profilo discriminatorio, il problema potrebbe sembrare irrisolvibile perché l'essere umano, dal quale gli algoritmi decisori traggono nutrimento, tende inevitabilmente a classificare il mondo che lo circonda secondo parametri istintivi e non razionali, o addirittura illegali o illeciti. Al contempo si pone la questione se sia ragionevole, ovvero eticamente appropriato, filtrare i contenuti dei database secondo una impostazione ideale della realtà, ingannando completamente la sua rappresentazione empirica.

Pertanto, l'implementazione di siffatti sistemi necessita di un approccio multidisciplinare, poiché la complessità dell'argomento trattato comporta il ricorso, essenziale, ad altri ambiti della conoscenza umana che possano supportare la scienza giuridica nella comprensione, e quindi nella disciplina, dei sistemi decisori automatizzati.

Per raffigurare il necessario scambio tra conoscenze in questo ambito si possono prendere in prestito le parole che Oscar Wilde scrisse nel suo libro più noto, "Il ritratto di Dorian Gray": *"Io sono geloso di ogni cosa la cui bellezza non muore. Son geloso del ritratto che mi hai dipinto, perché esso conserverà ciò che io devo perdere. Ogni momento che passa, a me ruba qualcosa e gliene aggiunge a lui. Oh! Se noi potessimo soltanto fare cambio! Se il quadro mutasse ed io rimanessi sempre così! Perché l'hai dipinto? Esso mi burlerà un giorno; si riderà di me orribilmente"*<sup>22</sup>.

Questo passaggio descrive in modo appropriato il rapporto tra privacy e discriminazione nell'elaborazione dei procedimenti di decisione automatizzata. Esso rappresenta una metafora di come la realtà appare, nascondendosi. Le parole di Wilde sembrano svelare il procedimento contenuto negli algoritmi decisori automatizzati, i quali assorbono le caratteristiche (tanto positive quanto negative) della persona raffigurata, mentre il passaggio antidiscriminatorio riguarderebbe la comprensione di cosa il ritratto stesso ruba al protagonista per aggiungerlo al risultato del procedimento automatizzato.

Tenendo a mente tale immagine, l'analisi e, successivamente, la ricostruzione, del rapporto tra privacy e discriminazione segue dapprima il percorso fattuale, e poi logico, della rivelazione di sé attraverso la diffusione dei propri dati personali.

Ciascuna persona, nel corso della sua vita, sparge, consapevolmente o inconsapevolmente, informazioni su di sé, le quali vengono raccolte, classificate ed elaborate da molteplici enti, che profilano i loro utenti per gli scopi più diversi (ad esempio: sicurezza, pubblicità, politica e così via). Siffatte operazioni vengono concluse al fine di predisporre strumenti predittivi, decisori ovvero manipolatori<sup>23</sup>, utili a

---

<sup>22</sup> Dal *Ritratto di Dorian Gray*, p. 45, e-book di liber liber, trad. it. di S. Ortolani.

<sup>23</sup> K. Crawford, *Né artificiale Né Intelligente. Il lato oscuro dell'IA*, trad. it. G. Arganese, Il Mulino, Bologna, 2021, *passim*.

giustificare una deliberazione in linea di massima già predisposta da valutazioni umane precedenti.

In altri termini, la decisione automatizzata rappresenta il bilanciamento ottimale tra l'efficienza nell'ottenimento di un risultato il più possibile aderente alla realtà seriale di riferimento e la necessità che il suo committente (singolo, ente pubblico o privato) ha di rimanere immune da forme di responsabilità rispetto alla decisione medesima, in particolare in termini di *accountability*, di fronte alla pubblica opinione.

Siffatta è l'idea che fonda questo libro e sulla quale si vuole analizzare il rapporto reciproco tra principio di non discriminazione, privacy e decisioni automatizzate. Durante il percorso di ricerca è maturata in chi scrive la convinzione che solo il controllo (soprattutto individuale e personale, sulla base del principio di autore-sponsabilità, oltre a quello istituzionale) dei propri dati possa rappresentare una soluzione il più possibile efficace contro il trattamento discriminatorio dei risultati dei procedimenti decisorii automatizzati. Tuttavia, rimane aperta la discussione su quale possa essere il rimedio effettivo con cui esercitare, sotto il profilo istituzionale, tale controllo, dato che in molte situazioni empiriche è diventato ormai impossibile sottrarsi alla tracciabilità dei propri dati.

A questo proposito, la disciplina europea in materia di privacy, in particolare il GDPR<sup>24</sup>, è volta a consentire a ciascun individuo di vincolare le piattaforme e gli operatori online alle proprie scelte in tema di cessione di dati personali utili alla profilazione degli utenti secondo le caratteristiche personali e sensibili. Tali dati sono dispersi dai loro soggetti-fonte sia nell'ambiente in cui vivono, tanto quello domestico quanto quello lavorativo, intercettati da strumenti di IoT (Internet of things), ovvero tramite l'uso delle piattaforme online (come Amazon, Facebook, Twitter e più di recente, il c.d. Metaverso), sistemi che di tali dati si nutrono.

Risulta pertanto rilevante analizzare come da un lato l'Europa, nello specifico l'Unione Europea, stia cercando di rafforzare la propria sovranità salvaguardando i dati, personali e non personali, dei propri cittadini nei confronti di piattaforme ed enti riferibili a Paesi terzi. In questo senso, il quadro giuridico formato dal GDPR (2018), dal Data Governance Act (2020), dall'Artificial Intelligence Digital Age Proposal (2021), Digital Market Service Act (2022) e dal Data Act Proposal (2022) intenderebbe recuperare sotto il profilo regolatorio il tempo perduto durante il decennio 2005-2015, periodo impiegato dalle c.d. GAFAM<sup>25</sup>

---

<sup>24</sup> Con questa sigla (acronimo di General Data Protection Regulation) ci si riferisce al Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati).

<sup>25</sup> Acronimo con cui vengono indicate le multinazionali statunitensi dell'informatica e delle piattaforme online Google (ora Alphabet), Apple, Facebook (ora Meta), Amazon e Microsoft (T. Stuart,

a diventare operatori dominanti sul mercato globale<sup>26</sup>.

Trova quindi interesse l'analisi di come il concetto di privacy viene inteso sulle due opposte sponde dell'Atlantico, in Europa e negli Stati Uniti, e come la privacy stessa viene posta in relazione alla questione antidiscriminatoria. Che cosa è veramente discriminatorio?

La risposta è solo apparentemente semplice. Essa non riguarda la ovvia definizione delle differenze, ma concerne il loro trattamento, specie se caratterizzanti l'identità della persona, come sesso, orientamento sessuale, identità etnica ovvero religiosa, capacità psicomotorie ovvero cognitive, inserite in un certo contesto. Tale contesto può rovesciare il trattamento discriminatorio, ovvero la sua percezione: essere biondi, per esempio, può essere considerato discriminatorio? Dipende. Infatti, è altresì importante verificare come viene effettuata la raccolta dei dati che alimentano gli ADMs, cioè algoritmi decisori automatizzati, e come la caratteristica in questione interagisce con l'insieme delle informazioni che compongono il database.

Questo approccio è interessante perché verifica come la raccolta dei dati che nutrono gli ADMs tende a raggruppare certi profili con altri profili simili ed è del tutto discutibile se sia legittimo ovvero eticamente giustificato (o meno) essere profilati secondo un certo gusto ovvero propensione all'acquisto, cioè le caratteristiche rilevanti per quel che riguarda il marketing o il diritto dei consumatori, ma non il diritto antidiscriminatorio.

Sul punto, occorre osservare che il diritto antidiscriminatorio non impedisce l'adozione di specifiche regole applicabili agli algoritmi decisori. Infatti, l'art. 22 GDPR<sup>27</sup> già prevede obblighi di trasparenza, spiegabilità ovvero revisione umana a carico delle decisioni algoritmiche. Tranne in questa fattispecie, ovvero in specifici casi previsti da norme ad hoc, non sussistono previsioni normative che mettano in dubbio, contestino ovvero annullino gli effetti giuridici attribuiti alle decisioni provenienti dagli ADMs, della loro validità ovvero applicabilità in caso di conseguenze giuridiche inerenti alla violazione del principio di non discriminazione da parte dell'algoritmo decisionario in quanto tale, ma deve essere applicata la normativa antidiscriminatoria<sup>28</sup>.

Alla luce di tali considerazioni ci si può chiedere quale sia lo strumento regolatorio più appropriato per la disciplina degli aspetti discriminatori legati all'utilizzo

---

*Too little too late? An exploration and analysis of the inadequacies of antitrust law when regulating GAFAM data-driven mergers and the potential legal remedies available in the age of Big Data*, in *European Competition Journal*, 2021, 17(2), p. 407 ss.).

<sup>26</sup> C. Gallese, *A First Commentary to the proposal for a new Regulation on fair access and use of Data (Data Act)*, in corso di pubblicazione, per gentile concessione dell'autrice.

<sup>27</sup> Si tratta dell'art. 22 rubricato "Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione" del Regolamento Generale sulla Protezione dei Dati n. 679/2016.

<sup>28</sup> T. Rodríguez de las Heras Ballell, *Guiding Principles for Automated Decision-Making in the EU*, European Law Institute, Vienna, 2022, p. 14 ss.

degli algoritmi decisori. In materia si sono fatti strada due modelli di regolamentazione, ovvero la predisposizione di una disciplina generale dell’“Artificial Intelligence”, come sembrerebbe essere orientata l’Unione europea, con l’Artificial Intelligence Act Proposal, oppure la previsione di norme specifiche all’uopo delineate, come accaduto in Canada con l’ADM Directive<sup>29</sup>, adottato a livello federale.

In merito alla prima ipotesi, il 21 aprile 2021, la Commissione europea e il Consiglio europeo hanno diffuso la Proposta di Regolamento in materia di Intelligenza Artificiale<sup>30</sup>, la quale si inserisce in un panorama regolamentare frammentato, considerando che le norme vigenti al momento più affini al tema riguardano la disciplina europea sulla protezione dei dati personali, in particolare l’art. 22 GDPR sul diritto ad accedere alla spiegazione della decisione automatizzata e la disciplina convenzionale, comunitaria e nazionale, sul divieto di discriminazione<sup>31</sup>.

In questo ambito è stata la giurisprudenza a supplire alle difficoltà del diritto positivo nell’elaborare soluzioni. La portata globale degli operatori che utilizzano strumenti di intelligenza artificiale ha consentito agli ordinamenti di esprimersi in ordine sparso, tuttavia raggiungendo, attraverso percorsi variegati, risultati omogenei, come si analizzerà in particolare nel quarto capitolo.

Da un punto di vista di policy, la Proposta di Regolamento UE in materia di intelligenza artificiale sembra avere quale obiettivo la replica dell’esperienza del GDPR, cioè imporsi come modello di regolamentazione di riferimento nell’ambito dell’Intelligenza artificiale come accaduto per il GDPR in materia di privacy<sup>32</sup>, soprattutto nella protezione transazionale dei dati personali raccolti massivamente, come il contenzioso Schrems ha dimostrato<sup>33</sup>.

In relazione alla specifica disciplina sugli ADM, le esigenze di regolamentazione concernono la gestione dell’opacità, complessità, dei bias, dell’imprevedibilità

<sup>29</sup> Directive on Automated Decision-Making, (2019) [Canada ADM Directive]. In dottrina, R. Wright, *Comparing European and Canadian AI Regulation*, Law Commission of Ontario, Toronto, 2021, *passim*; T. Scassa, *Administrative Law and the Governance of Automated Decision Making: A Critical Look at Canada’s Directive On Automated Decision Making*, in *UBCL Rev.* 54, 2021, p. 251.

<sup>30</sup> Proposta di Regolamento del Parlamento Europeo, *cit.*

<sup>31</sup> J. Della Torre, *Le decisioni algoritmiche al vaglio del Consiglio di Stato*, in *Riv. Dir. Proc.*, 2021, 2, p. 710 ss.

<sup>32</sup> J. Kessler, *Data Protection in the Wake of the GDPR: California’s Solution for Protecting “the World’s Most Valuable Resource”*, in 93 *S. Cal. L. Rev.* 99 (2019-2020), p. 111 ss.; E.L. Field, *United States Data Privacy Law: The Domino Effect After the GDPR*, in 24 *N.C. BANKING INST.* 481 (2020), p. 483 ss.; M.L. Rustad, T.H. Koenig, *Towards a Global Data Privacy Standard*, in 71 *Fla. L. Rev.* 365 (2019), p. 411 ss.

<sup>33</sup> F. Bignami, G. Resta, *Transatlantic Privacy Regulation: Conflict and Cooperation*, in 78 *Law and Contemporary Problems*, 2015, pp. 231-266; A. Chander, *Is Data Localization a Solution for Schrems II?*, in *Journal of International Economic Law*, vol. 23, Iss. 3, September 2020, pp. 771-784; H.J. Brehmer, *Data Localization The Unintended Consequences Of Privacy Litigation*, in *American University Law Review*, 2018, vol. 67: Iss. 3, Article 6; S. Dighera, M. Giraud, *Privacy rights in online interactions and litigation dynamics: A social custom view*, in *European Journal of Political Economy*, 2021.

del comportamento parzialmente autonomo dei certi sistemi automatizzati al fine di assicurarne la compatibilità con la protezione dei diritti fondamentali e l'applicazione delle norme giuridiche<sup>34</sup>.

La Proposta sembra aver assorbito l'esperienza empirica predisponendo la disciplina la materia secondo gli orientamenti giurisprudenziali affermatasi nelle corti, specie in materia di decisioni automatizzate nei confronti delle persone vulnerabili<sup>35</sup>, ovvero relativamente sulla profilazione effettuata dagli algoritmi in ambito lavorativo<sup>36</sup>. È in questo senso che si propone di classificare gli algoritmi automatizzati secondo la loro "classe di rischio" ed i principi di proporzionalità e necessità, qualificandone il livello e consentendone l'utilizzo soltanto in adesione a parametri rigidi predeterminati o addirittura proibendoli<sup>37</sup>.

Il quadro legale che ne deriverebbe seguirebbe un "*risk-based approach*" nell'imposizione di limiti all'utilizzo di *high risk AI systems* solo quando questi obbediscano a requisiti certi quali l'utilizzo di dati di alta qualità, tracciati documentati e trasparenti, sotto il controllo di referenti umani che ne certifichino l'accuratezza e la robustezza. Secondo la Commissione Europea tali parametri sono strettamente necessari per mitigare il rischio di violazione di diritti fondamentali e dei parametri di sicurezza posto in essere dall'uso dei sistemi intelligenti e che non sono, al momento, protetti da alcuna garanzia giuridica coordinata, dato che l'approccio basato sulla sola legislazione nazionale provocherebbe un patchwork normativo foriero di incertezze giuridiche e barriere alla circolazione dei prodotti e dei servizi.

Analizzando tale Proposta di Regolamento sembrerebbe sensato affermare che esso è focalizzato sulla protezione dei diritti fondamentali, dato che impone restrizioni sulle libertà di ricerca scientifica e di impresa (per esempio proibendo l'uso di certune tecnologie, come il riconoscimento biometrico istantaneo ovvero di *social scoring*). Il fine di tali restrizioni è di assicurare che nel bilanciamento tra interessi protetti prevalgano le ragioni della salute, della sicurezza, della protezione dei consumatori, della non discriminazione negli ambiti dove le tecnologie di intelligenza artificiale ad alto rischio sono sviluppate e impiegate. Tali limitazioni devono essere proporzionate e ridotte al minimo necessario per prevenire e mitigare i rischi plausibili di violazione dei diritti umani protetti dalla Carta dei diritti umani dell'Unione europea. Tra questi si ricordano innanzitutto la protezione della dignità umana (art. 1) attraverso il rispetto della vita privata (art. 7), della protezione dei dati personali (art. 8), del principio di non discriminazione (art. 21), dell'uguaglianza tra uomini e donne (art. 23), che devono essere estesi a coloro che non si ricono-

---

<sup>34</sup> Proposta di Regolamento del Parlamento Europeo, cit., p. 3, *Council of the European Union, Presidency conclusions – The Charter of Fundamental Rights in the context of Artificial Intelligence and Digital Change*, 11481/20, 2020.

<sup>35</sup> Ci si riferisce alla decisione olandese SyRI, *supra*, Cap. IV.

<sup>36</sup> Si veda il contenzioso che contrappone i riders e i drivers contro le piattaforme di gig-economy come Uber o Foodora, *supra*, Cap. IV.

<sup>37</sup> Proposta di Regolamento del Parlamento Europeo, cit.

scono in una categoria di genere binaria. Ulteriormente, devono essere prevenuti i rischi che l'uso di programmi di intelligenza artificiale possono provocare sulla libertà di espressione (art. 11) e di assemblea (art. 12)<sup>38</sup>, assicurare il diritto all'esercizio di un rimedio effettivo e a un giusto processo con il riconoscimento del diritto alla difesa e a un giusto processo art. 43 e art. 48). Altresì, le decisioni automatizzate effettuate con algoritmi possono avere effetti negativi su un indefinito numero di gruppi di soggetti vulnerabili protetti dalla Carta dei diritti fondamentali dell'Unione Europea, come i disabili (art. 26), lavoratori (art. 31), minori (art. 24) o minoranze. La protezione della salute e della sicurezza dei consociati può essere raggiunta anche attraverso un alto livello di protezione ambientale (art. 37).

Il rispetto di siffatti obblighi giuridici deve essere attuato attraverso test preliminari dei programmi automatizzati affinché siffatti test di utilizzo ne minimizzino i rischi di errori o di risultati falsati dal bias, soprattutto in aree critiche quali quelle relative ai servizi sociali, di educazione, lavoro e giudiziari. Va osservato che si parla di minimizzazione, ma non di eliminazione in senso assoluto per la modalità di funzionamento del *machine learning* alla base di ogni algoritmo automatizzato<sup>39</sup>. Se la violazione dei diritti fondamentali permane, il Proposal di regolamento richiede un "*effective redress*" per le persone coinvolte attraverso trasparenza e tracciabilità dei sistemi automatizzati da accompagnarsi con controlli *ex post*.

Lo scopo delle restrizioni imposte dalla Proposta della Commissione europea alle libertà di impresa (art. 16 della Carta) e alla libertà di ricerca scientifica e tecnologica (art. 13) è di attuare una "innovazione responsabile" al fine di bilanciamento con le superiori ragioni di tutela degli interessi pubblici (quali salute, sicurezza, protezione dei consumatori) e i summenzionati diritti fondamentali. Ci si chiede cosa ciò possa significare e a questo proposito è la Proposta stessa a prefigurare tale responsabilità nel mitigare i rischi seri di sicurezza e le probabili violazioni dei diritti umani.

La Proposta affronta anche il tema della tutela della proprietà intellettuale, di tutela del know how o di informazioni confidenziali sui beni immateriali (ex art. 17.2 della Carta), la cui protezione non deve inficiare il principio di trasparenza, affinché le parti colpite dagli effetti avversi degli algoritmi siano poste in grado di esercitare il loro diritto a un rimedio effettivo e alla trasparenza informativa. Siffatto problema si è già presentato in materia di accesso ai codici software dei programmi di decisione automatizzata relativamente all'applicazione dell'art. 22 GDPR sul diritto all'accesso alle informazioni relative alle decisioni automatizzate.

Secondo l'Allegato III della Proposta, i programmi di intelligenza artificiale a rischio inaccettabile sono il riconoscimento facciale biometrico e il *social scoring*:

---

<sup>38</sup> In materia di possibile limitazione alla libertà di manifestazione del pensiero e di assemblea attuata attraverso algoritmi biometrici, si ricorda il caso *Leaders of a Beautiful Struggle v. Baltimore Police Department* in cui la *United District Court for the District of Maryland* ha deciso un rilevante caso in materia di privacy e sicurezza svoltosi a Baltimora discusso nel Cap. IV, par. 3.

<sup>39</sup> *Supra*, Cap. III.

il primo però viene autorizzato in casi specifici che in realtà ne consentono un uso potenzialmente molto frequente, seppur a seguito del controllo giudiziario; il secondo si riferisce a pratiche che difficilmente possono venire attuate in ordinamenti che riconoscono la *rule of law*, ma in regimi autoritari<sup>40</sup>.

Il *social scoring* concerne programmi di intelligenza artificiale ad alto rischio, soggetti all'adeguamento con specifici requisiti e conformità alla normativa valutabile *ex ante*, come i software di profilazione utilizzati in nell'assunzione dei lavoratori, nell'accesso ai benefici sociali, o in caso di uso di *medical devices*. Vi sono poi algoritmi che sono permessi ma solo se soggetti a obblighi di informazione e trasparenza (*bots*), mentre i programmi con rischi minimi o del tutto assenti sono consentiti senza restrizioni (sistemi esperti).

Le critiche che sono state dirette contro tale proposta riguardano la vaghezza dei concetti, in particolare la definizione di “*adverse impact*” (contenuto nell’art. 7 della Proposta), la formula “*AI system intended to be used*” utilizzata dall’Allegato III garantisce una interpretazione molto ampia, ai limiti dell’analogia e le criticità principali nell’uso del riconoscimento biometrico “*real time*” ovvero “*post*” (da remoto) la cui asseverazione della conformità alla normativa viene ceduta a terze parti, creando potenziali conflitti di interessi<sup>41</sup>, soprattutto al riguardo di quali sistemi biometrici debbano essere destinati a tale controllo, cioè quelli di *predictive policing*, controllo della migrazione e *recruitment*.

Ulteriormente, la possibilità stessa di distinguere tra le categorie di programmi di intelligenza artificiale ad alto rischio e quelle a basso rischio è scorretta perché vi sono, software che interagiscono con le persone, come i *chatbots*, che possono infliggere pesanti conseguenze agli utenti. A questo proposito si segnala l’esperienza dei sistemi automatici di risposta utilizzati dal programma welfare pubblico dell’Universal Credit britannico<sup>42</sup> ovvero i programmi di riconoscimento emozionale<sup>43</sup> o quelli in grado i sistemi automatizzati in grado di produrre *deepfakes* manipolando i contenuti. In altri termini, sistemi di riconoscimento emotivo e di categorizzazione biometrica, nonché quelli che coinvolgono i *deepfakes*<sup>44</sup>, sono applica-

---

<sup>40</sup> N. Maamar, *Social Scoring: Eine europäische Perspektive auf Verbraucher-Scores zwischen Big Data und Big Brother*, in *Computer und Recht*, vol. 34, n. 12, 2018, pp. 820-828. <https://doi.org/10.9785/cr-2018-341212>; N. Cristianini, T. Scantamburlo, *On social machines for algorithmic regulation*, in *AI & Society*, 35, pp. 645-662, 2020; <https://doi.org/10.1007/s00146-019-00917-8>.

<sup>41</sup> F. Reinhold, A. Müller, *AlgorithmWatch’s response to the European Commission’s proposed regulation on Artificial Intelligence – A major step with major gaps, 22.4.201*, <https://algorithmwatch.org/en/response-to-eu-ai-regulation-proposal-2021/>.

<sup>42</sup> *Supra*, Cap. IV.

<sup>43</sup> F.Z. Canal, T.R. Müller, J.C. Matias, G.G. Scotton, A.R. de Sa Junior, E. Pozzebon, A.C. Sobieranski, *A survey on facial emotion recognition techniques: A state-of-the-art literature review*, in *Information Sciences*, 2022, pp. 582, 593-617.

<sup>44</sup> M. Groh, Z. Epstein, C. Firestone, R. Picard, *Deepfake detection by human crowds, machines, and machine-informed crowds. Proceedings of the National Academy of Sciences*, 2022, 119(1), e21100131119; L.A. Passos, D. Jodas, K.A. da Costa, L.A.S. Júnior, D. Colombo, J.P. Papa,

zioni che possono presentare un alto potenziale di grave danno per gli individui e le società democratiche. Inoltre, la base scientifica dei sistemi di riconoscimento, soprattutto emotivo, è molto controversa. Gli obblighi di trasparenza già minimi che si applicano a questo insieme di sistemi di IA sono ulteriormente indeboliti da una serie di esenzioni (come per l'uso di *chatbot* o sistemi di categorizzazione biometrica per la prevenzione, l'indagine o il perseguimento dei crimini)<sup>45</sup>.

Per quanto concerne la regolamentazione settoriale degli ADM, la summenzionata Direttiva canadese, già in vigore dal 1 aprile 2020, riguarda invece l'adozione di specifici principi giuridici applicabili in ambito amministrativo federale nell'erogazione di servizi esterni<sup>46</sup>. Tale circostanza limita sensibilmente gli effetti della disciplina, dato che essa non è applicabile né all'utilizzo dei sistemi decisori automatizzati in ambito pubblicistico relativamente ai governi ovvero agenzie provinciali, come polizia locale e i servizi sociali, né in ambito privatistico, vanificando un possibile positivo impatto antidiscriminatorio sulla società nel suo complesso.

Per quanto riguarda il modello di disciplina settoriale, sebbene l'*ADM Directive* vincoli in via residuale soltanto parte dell'apparato federale pubblico canadese, permane di interesse riflettere su come essa si relazioni sugli aspetti critici dell'uso di tali sistemi, in particolare per quel che concerne la valutazione di livelli di rischio e sull'impatto degli effetti dell'uso degli ADMs sugli utenti. Siffatta valutazione viene svolta attraverso il procedimento previsto dall'*Algorithmic Impact Assessment*<sup>47</sup> effettuato attraverso l'indagine basata su un questionario composto da 60 domande sulle caratteristiche del sistema: per esempio, in relazione alle modalità della raccolta dati, su quale tipo di algoritmo sia stato utilizzato, se il medesimo preveda un codice sia aperto ovvero sia protetto da diritti di proprietà intellettuale, se i soggetti destinatari del risultato algoritmico appartengano a categorie vulnerabili e così via. Tale metodo non sembra efficiente perché riduce l'analisi dell'impatto algoritmico ad un fatto meramente descrittivo focalizzato sull'utilizzo del singolo utente. Al contrario, la trasparenza, e quindi la comprensione, dell'algoritmo decisionario emergono nel momento in cui il funzionamento e i risultati ottenuti sono confutabili ed analizzabili empiricamente su un grande numero di casi, e quindi di dati, trattati.

Ciò nonostante, la criticità più importante di tale disciplina riguarda la circostanza che esplicitamente essa non vada applicata in fattispecie nelle quali l'uso di algoritmi decisori comporti il verificarsi di discriminazioni, come nei "*Systems (that)*

---

*A Review of Deep Learning-based Approaches for Deepfake Content Detection*, 2022, arXiv preprint arXiv:2202.06095.

<sup>45</sup> F. Reinhold, A. Müller, *Algorithm Watch's response*, Cit.; J. Wulf, *Automated Decision-Making Systems and Discrimination. Understanding causes, recognizing cases, supporting those affected*, AW AlgorithmWatch gGmbH, Berlin, 2022, *passim*.

<sup>46</sup> Directive on Automated Decision Making, § 4.2.2 e 4.2.3. In dottrina, T. Scassa, *op. cit.*

<sup>47</sup> Directive on Automated Decision Making, Annex C., T. Scassa, *op. cit.*, R. Wright, *Comparing European and Canadian AI Regulation*, cit.

*could be deployed in the criminal justice system or criminal proceedings*”<sup>48</sup>. In considerazione della duplice limitatezza della applicabilità della disciplina canadese sia dal punto di vista formale, essendo essa confinata a livello federale, sia la ristrettezza delle materie cui è rivolta, sembrerebbe che siffatta esperienza che non sia in grado di apportare validi contributi costruttivi.

In conclusione, l’*ADM Directive* è una normativa settoriale più vicina al *soft law* che all’*hard law*, e nonostante riguardi la competenza federale dell’ordinamento canadese, si rivela essere insufficiente nel trattare il rischio discriminatorio connesso all’uso massivo di algoritmi decisori automatizzati in quanto inefficace nel disciplinare le moltitudini di fattispecie potenzialmente discriminatorie gestite con tali sistemi.

Tale convinzione è sorta in chi scrive a seguito dell’analisi empirica proposta nel primo capitolo di questo testo, in relazione alla radicata presenza della discriminazione nella società umana, tanto in contesti storici e sociali ormai trascorsi, ma il cui retaggio rimane presente e invadente, quanto in quella attuale in cui gli algoritmi operano.

Il principale rimedio antidiscriminatorio proposto riguarda la tutela della privacy quale schermo protettivo dalle indagini invasive sui big data trattati dagli algoritmi. A questo proposito, nel secondo capitolo ci si focalizza su tale tutela, relativamente all’utilizzo dei dati diffusi dagli utenti, sia nel caso siffatto utilizzo abbia almeno un margine di consapevolezza del titolare dei dati raccolti, come ad esempio attraverso le reti sociali, ovvero inconsapevole come accade ai dati raccolti per mezzo della c.d. “Data Retention”. Altresì viene analizzato come questi dati, una volta raccolti, vivano di una vita propria della quale la persona a si riferiscono ignora l’esistenza. Infatti, i dati rappresentano le loro fonti originarie indipendentemente dalla volontà o consapevolezza di queste ultime, come accaduto nelle vicende relativa a “Cambridge Analytica”, o ai “Facebook Files”, dove la profilazione di ignari utenti dei social network è servita per orientare l’opinione pubblica nelle consultazioni elettorali<sup>49</sup>.

Di fronte a questo scenario gli ordinamenti appartenenti alla Western Legal Tradition hanno reagito in modo opposto: da un lato negli Stati Uniti attraverso la regolamentazione dell’accesso quasi indiscriminato ai dati personali e sensibili dei consociati, motivando questa “trasparenza di Stato” con ragioni di sicurezza nazionale, mentre dall’altro lato l’Unione Europea ha implementato una rigorosa protezione della privacy attraverso il Regolamento UE 679/2016. Tale contrapposizione ha generato una lunga controversia di fronte alla Corte di Giustizia dell’Unione Europea (e di fronte alle autorità nazionali, in particolare quella irlandese) sul trasfe-

---

<sup>48</sup> R. Wright, cit. *Supra*, Cap. IV.

<sup>49</sup> L. Rizzo, *Harvesting your soul? Cambridge analytica and Brexit*, in C. Jansohn (ed.), *Brexit Means Brexit*, Mainz, Academy of Sciences and Literature, 2018, pp. 75-90; M. Hu, *Cambridge Analytica’s black box*, in *Big Data & Society*, 7(2), 2020, 2053951720938091.

rimento dei dati personali dei cittadini europei utenti di piattaforme statunitensi e quali siano stati i rimedi, principalmente giudiziari (si vedano sia i contenziosi “Google Spain”<sup>50</sup>, sia Schrems I<sup>51</sup> e Schrems II<sup>52</sup>) per un tempo che in realtà è indefinito e che pertanto rischia di diventare definitivo perché senza limiti (cioè per sempre).

Si è data particolare attenzione a questo contenzioso, alla protezione della privacy e dei dati personali per analizzare e comprendere come la normativa a tutela dei dati personali possa organizzare la produzione del carburante, cioè le informazioni massive, delle scatole nere (black box) produttive di decisioni automatizzate e, in quanto automatizzate, potenzialmente discriminatorie.

Il terzo capitolo è concentrato sulla comprensione di cosa sia una *black box*, la scatola nera cuore del *machine learning*, di come quest’ultimo funzioni coordinandosi con le c.d. “reti neurali” e di come siffatta combinazione rappresenti una imitazione delle sinapsi che caratterizzano il ragionamento umano al fine di comprendere quanto una decisione automatizzata sia imputabile all’algoritmo che l’ha elaborata oppure al soggetto umano che quell’algoritmo ha creato e nutrito con delle informazioni precedentemente selezionate o orientate in un certo modo. Del pari, si approfondisce cosa sia e come funziona il diritto di spiegazione della decisione automatizzata, cioè lo strumento giuridico che la persona sottoposta a decisione elaborata dall’algoritmo dispone per verificare se ha subito un trattamento discriminatorio, o meno, da parte del decisore automatico. Questo capitolo è stato elaborato tenendo presente quale sia la natura dell’algoritmo, quali siano le condizioni di trasparenza del procedimento decisorio, come stabilito empiricamente dai giudici nazionali che hanno applicato la normativa europea in materia, cioè l’art. 22 del GDPR.

Il quarto capitolo è dedicato all’analisi nel dettaglio del rapporto tra bias, cioè il fattore qualificante di un certo comportamento o di una caratteristica identitaria, e la discriminazione, attraverso lo studio di specifici casi empirici, quali l’utilizzo dei *risk assessment algorithms* nella raccolta dei dati biometrici, nell’accesso agli studi universitari, nell’organizzazione delle mansioni lavorative e nell’erogazione dei benefici sociali. Siffatte non sono le uniche situazioni dove possono venire utilizzati gli algoritmi decisori automatizzati, ma sono quelle in cui l’aspetto discriminatorio nella decisione delegata a un sistema automatizzato può avere conseguenze nocive per le parti discriminate coinvolte con lesione dei loro diritti fondamentali.

Al termine di questa parte introduttiva si intende specificare che nel testo sovente è stata utilizzata la locuzione “giudice umano” in riferimento alle decisioni in cui

---

<sup>50</sup> Corte di Giustizia dell’Unione Europea, (Grande Sezione), 13 maggio 2014, C-131/12, Google Spain SL e Google Inc. contro Agencia Española de Protección de Datos (AEPD) e Mario Costeja González.

<sup>51</sup> Corte di Giustizia dell’Unione Europea, (Grande Sezione), 6 ottobre 2015, C-362/14, Maximilian Schrems contro Data Protection Commissioner.

<sup>52</sup> Corte di Giustizia dell’Unione Europea, (Grande Sezione), 16 luglio 2020, C-311/18, Data Protection Commissioner contro Facebook Ireland Limited e Maximilian Schrems.

necessariamente è presente l'*humanitas* decisoria contrapponendola all'algoritmo automatizzato. L'uso di tale termine è utile per sottolineare l'imprescindibile capacità di discernere, tipica degli esseri umani, ma non degli algoritmi.

Infine, premettendo che tutti gli errori e le omissioni di questo saggio sono propri ed esclusivi dell'autrice, chi scrive desidera ringraziare coloro che si sono resi disponibili al confronto intellettuale e al sostegno del progetto che ha condotto all'elaborazione di questo lavoro.

## CAPITOLO PRIMO

# UGUAGLIANZA, DISUGUAGLIANZA E DISCRIMINAZIONE

Sommario: 1. Diseguale e discriminatorio: le radici di un problema apparentemente senza soluzione. – 2. Rivoluzioni dei diritti e rivoluzione industriale: l'avvento dell'automazione. – 3. Il retaggio della schiavitù nel contrasto a disuguaglianza e non discriminazione. – 4. L'elaborazione di una strategia antidiscriminatoria e il ruolo degli stereotipi. – 5. Scelte tragiche e applicazione empirica del principio di non discriminazione. – 6. La difficoltà di individuare e di sanzionare una discriminazione: il contrasto dell'hate speech. – 7. Esposizione della personalità digitale e conseguenze discriminatorie nell'automatizzazione delle decisioni.

### 1. Diseguale e discriminatorio: le radici di un problema apparentemente senza soluzione

Nella tradizione occidentale, fino al declino dell'Età moderna, il concetto di uguaglianza apparteneva, più che a questo, all'altro mondo, cioè a quello dei già vissuti, in attesa del Giudizio Universale dove tutti sarebbero stati ugualmente giudicati per gli atti commessi in vita<sup>1</sup>. Le società passate, infatti, si distinguevano in base a status acquisiti per nascita di liberi, servi o schiavi. Tali distinzioni potevano essere, in certi casi, basate sul diritto divino<sup>2</sup> ovvero sull'ordine costituito<sup>3</sup>.

Il concetto di uguaglianza rappresentativo di persone uguali per nascita, diritti e desideri è più recente e prende forma con l'Illuminismo. La filosofia morale di Kant consente di elaborare l'idea di uguaglianza nella prospettiva della dignità: la dignità è riconosciuta in quanto vi è uguaglianza tra i consociati. Infatti i tre principi della società civile sono la libertà di ciascun membro della società in quanto persona, l'indipendenza di ciascun in quanto cittadino e l'uguaglianza di ciascuno con gli altri in quanto appartenente ad una comunità<sup>4</sup>, i cui membri sottoposti alle me-

---

<sup>1</sup> B.D. Ehrman, *Heaven and Hell. A History of the Afterlife*, Simon and Schuster, New York, 2020, *passim*.

<sup>2</sup> M. Bloch, *La società feudale*, trad. it. B.M. Cremonesi, Einaudi, Torino, 1999, *passim*.

<sup>3</sup> P. Gallo, voce "Soggetto di diritto", *Digesto Civile*, UTET, Torino, 1998, Vol. XVIII, p. 576 ss.

<sup>4</sup> I. Kant, *Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die*

desime leggi, dato che ogni membro della società civile ha nei confronti degli altri degli obblighi (diritti coattivi)<sup>5</sup>.

L'uguaglianza intesa in senso giuridico non è un'idea vuota<sup>6</sup>, ma un concetto autonomo, complesso e plurale, anche da un punto di vista semantico, dato che dall'"uguaglianza"<sup>7</sup> può essere distinta l'"eguaglianza"<sup>8</sup>, termini spesso utilizzati come sinonimi. Infatti, il primo termine si riferisce ad una "(C)ondizione di cose o persone che siano tra loro uguali, cioè abbiano le stesse qualità, gli stessi attributi (in generale, o in un certo ambito)" mentre il secondo riguarda "la condizione per cui ogni individuo o collettività deve essere considerato alla stregua di tutti gli altri, e cioè pari, soprattutto nei diritti civili, politici, sociali ed economici. L'eguaglianza di tutti davanti alla legge è, assieme alla libertà, un diritto fondamentale dell'uomo e una delle regole-base di una convivenza democratica". Se il primo termine sembra riferirsi al concetto più in generale, mentre il secondo attiene ad un ambito specificamente giuridico, nonostante esso sembri trarre linfa dalle radici rivoluzionarie francesi del termine "égalité". Siffatte origini non nascondono la vana illusione di fissare il significato di uguaglianza, nonostante essa si riveli essere sfuggente. Infatti, già durante i Lumi venne elaborata la distinzione tra disuguaglianza fondata su cause naturali o fisiche, come condizioni di età, salute, caratteristiche fisiche ovvero spirituali e disuguaglianza basata su circostanze politiche o morali, riconducibili a contesti legati a condizioni sociali conseguenti alla ricchezza, al prestigio o all'onore che inducono altri consociati sprovvisti di tali caratteristiche all'obbedienza<sup>9</sup>.

Prendendo ora in prestito temporaneo un concetto che verrà utilizzato in prosieguo, cioè il *bias*, è possibile riflettere come esso possa contenere in sé gli elementi in grado di manifestare tratti discriminatori nei risultati del procedimento decisionario automatizzato. Pertanto è possibile anticipare, seppur sommariamente, che una decisione automatizzata effettuata attraverso un algoritmo non è necessariamente discriminatoria in sé, ma più probabilmente mantiene, incapace di eliminarle auto-

---

*Praxis*, 1793. Sul punto si veda M.A. Cattaneo, voce "Uguaglianza", Digesto UTET, Sezione pubblicistica, Torino, 1999, Vol. XV, p. 417 ss.

<sup>5</sup> M.A. Cattaneo, voce "Uguaglianza", cit.

<sup>6</sup> Quest'idea è stata promossa da diversi autorevoli autori, tra cui primeggia in lingua italiana N. Bobbio, secondo cui uguaglianza, e libertà, sono contenitori generici e vuoti che "se non specificati o riempiti non significano nulla" (N. Bobbio, *Eguaglianza ed egualitarismo*, in AA.VV., *Eguaglianza ed egualitarismo*, Armando Editore, Roma, 1978, p. 14. In lingua inglese si ricorda il contributo di P. Westen, *The Empty Idea of Equality*, in *Harvard Law Review*, 1982, pp. 537-596. Tratta diffusamente in maniera critica del concetto promosso da questo autore, L. Giacomelli, *Ripensare l'uguaglianza. Effetti collaterali della tutela antidiscriminatoria*, Giappichelli, Torino, 2018, p. 8 ss.).

<sup>7</sup> Voce "Uguaglianza", Vocabolario Treccani: <https://www.treccani.it/vocabolario/uguaglianza/>.

<sup>8</sup> Voce "Eguaglianza", Enciclopedia Treccani: <https://www.treccani.it/enciclopedia/eguaglianza>.

<sup>9</sup> È notorio che il fautore di tale classificazione fosse J.J. Rousseau, il quale nel *Discours sur l'origine de l'inégalité parmi les hommes*, del 1754 fondava la sua elaborazione teorica partendo dalla critica al concetto di proprietà privata (M.A. Cattaneo, voce "Uguaglianza", cit.).

mamente, le tendenze discriminatorie già presenti nella realtà.

Secondo il diritto naturale, come scriveva il Cavalier Louis de Jaucourt, l'«*Égalité naturelle (Droit nat.) est celle qui est entre tous les hommes par la constitution de leur nature seulement. Cette égalité est le principe et le fondement de la liberté. L'égalité naturelle ou morale est donc fondée sur la constitution de la nature humaine commune à tous les hommes, qui naissent, croissent, subsistent et meurent de la même manière*»<sup>10</sup>.

Tuttavia, in questa sede, si preferisce utilizzare il concetto di “uguaglianza”, in senso più ampio, non esclusivamente giuridico, dato che l'obiettivo di questo studio è analizzare le relazioni tra uguaglianza, discriminazione e l'uso degli algoritmi decisori nei diversi ambiti della vita dove tali procedimenti automatizzati vengono adoperati.

L'uguaglianza di fronte alla legge, in senso kantiano, si deve confrontare con le disuguaglianze involontarie che, collegate a circostanze fattuali, richiedono una speciale protezione a causa della loro condizione non voluta. Certa dottrina propone di classificare analoghe situazioni in gruppi di “microdiseguaglianze”<sup>11</sup>. Tale atteggiamento illustra una visione contemporanea della questione.

Dal punto di vista filosofico originalista<sup>12</sup>, l'uguaglianza che più urge realizzare riguarda quella giuridica, cioè inerente alla parità di trattamento di tutti i cittadini di fronte alla legge nel godimento dei diritti individuali<sup>13</sup>, in riferimento alla “comune umanità” delineata da Jacourt nella *Encyclopédie*. L'orientamento originalista affonda le sue radici nell'idealismo inglese del XVIII Secolo, che si riferisce ad un concetto di Stato neutrale e liberale, “equidistante dai conflitti sociali, culturali e religiosi, per il quale il diritto è oggettivo e caratterizzato da categorie astratte e generali”<sup>14</sup>. Per quanto qui di interesse, esso permea a tutt'oggi la mentalità giuridica americana<sup>15</sup> e ha un significativo impatto sui temi oggetto della nostra attenzione.

<sup>10</sup> L. de Jaucourt, «*Egalité naturelle*», *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*, Denis Diderot, Jean Le Rond d'Alembert (dir.), tome 5, Chez Briasson, David, Le Breton et Faulche (ed.), 1755, p. 415.

<sup>11</sup> M. Ainis, *La piccola eguaglianza*, Einaudi, Torino, 2015, p. 33 ss.

<sup>12</sup> Con “originalismo” si intende, specie nell'ambito del Common Law statunitense, l'obbedienza “to the original public meaning of today's Constitution drawn from the understanding and interpretation of the generation that ratified the Constitution of 1789” (M.S. Lewis, *Evil History: Protecting Our Constitution Through an Anti-Originalism Canon of Constitutional Interpretation*, in 18 *U.N.H. L. Rev.* 261 (2020), p. 261; C. Teerberk, “*Clocks Must Always Be Turned Back*”: *Brown v. Board of Education and the Racial Origins of Constitutional Originalism*, in *American Political Science Review*, 115(3), 2021, pp. 821-834).

<sup>13</sup> M.A. Cattaneo, *op. cit.*

<sup>14</sup> L. Giacomelli, *Ripensare l'uguaglianza*, cit., p. 73.

<sup>15</sup> La dottrina originalista americana trova un suo spazio fondativo nell'affermazione che devono essere i corpi elettorali, attraverso i loro rappresentanti, a decidere per mezzo di un atto legislativo la disciplina normativa delle fattispecie e non attraverso gli interventi (più o meno creativi) della giurisprudenza. L'espressione più recente e rigorosa di tale orientamento di pensiero si trova nella decisione *Dobbs v. Jackson Women's Health Organization* che ha effettuato l'overruling della notissima sen-

Apparentemente questa prospettiva, così familiare da sembrare semplice e ovvia, è stata assorbita negli ordinamenti costituzionali moderni, i quali hanno mantenuto una chiara matrice giusnaturalistica, al pari delle Carte dei diritti fondamentali che stanno alla base dello Stato di diritto<sup>16</sup> conosciuto finora. Lo spirito illuministico si identifica nelle Dichiarazioni dei diritti dell'uomo e del cittadino (1789) o il Bill of Rights americano (1791) emanate a partire dal Settecento e nelle costituzioni coeve. Infatti, la proclamazione del principio di uguaglianza nelle Carte costituzionali elaborate durante la Rivoluzione Francese “aveva il preciso intento di reagire e abbattere un ordinamento (l'Ancien Régime) fatto di privilegi divenuti, ormai, intollerabili”<sup>17</sup>.

La società sarebbe stata riorganizzata sulla base dell'uguaglianza. Essa rappresentava la pietra angolare sulla quale fondare “i diritti umani (...) e riformare sulle giuste basi la politica, le istituzioni, le relazioni sociali, il matrimonio, l'istruzione e le leggi”<sup>18</sup>.

Vi è un aspetto ulteriore di rilevanza sociale: è vero che l'uguaglianza dei singoli realizzava l'abolizione del Vecchio Regime feudale, il quale rappresentava un “cambiamento sconvolgente” per l'Europa continentale da secoli dominata dalla nobiltà<sup>19</sup>, cioè da chi vantava uno status privilegiato per nascita; ma l'incombente Rivoluzione Industriale avrebbe parcellizzato le categorie nella società di allora proponendo uno schema comunque legato alla disuguaglianza sociale, conseguente all'implementazione della tecnologia e alla redistribuzione della ricchezza.

In conseguenza di ciò apparirebbe che l'attuazione, almeno in via teorica, delle idee rivoluzionarie avesse comportato la sostituzione delle disuguaglianze legate agli status con le discriminazioni fattuali, a causa degli effetti dell'uso su scala sempre più larga delle innovazioni tecnologiche legate all'automazione applicata alla nascente industria. Nondimeno, questo ciclo non sembra affatto essersi interrotto, ma assurge nuova linfa vitale da ogni passaggio evolutivo della tecnologia.

Esiste quindi una comune matrice illuminista che promuove tre importanti Rivoluzioni: due politiche, cioè quella Americana e quella Francese, e una sociale, quella Industriale. Ci si può quindi chiedere quale sia il rapporto tra i tre avvenimenti, se vi è una correlazione tra le rivoluzioni politiche e quella sociale e se questa è ripetibile ora.

---

tenza *Roe v. Wade* in materia di aborto sostenendo che la Costituzione americana e il Bill of Rights vanno interpretati alla luce delle intenzioni dei Founding Fathers, cioè dei Padri Fondatori che li hanno elaborati, alla fine del XVIII secolo.

<sup>16</sup> M.A. Cattaneo, *op. cit.*

<sup>17</sup> M. Losana, “*Stranieri*” e principio costituzionale di uguaglianza, in *Rivista AIC*, 2016, 1, p. 11; F. Sorrentino, *Eguaglianza*, Giappichelli, Torino, 2011.

<sup>18</sup> J. Israel, *La Rivoluzione francese. Una storia intellettuale dai Diritti dell'uomo a Robespierre*, trad. it. P. Di Nunno, M. Nani, Einaudi, Torino, p. 30.

<sup>19</sup> J. Israel, *La Rivoluzione francese*, cit., p. 130.

Nei neonati Stati Uniti, apparentemente gli ex colonizzatori e gli ex coloni intendevano in modo antitetico l'idea di uguaglianza. Le ex colonie britanniche si differenziavano sotto il profilo della dis-uguaglianza tra i loro consociati (o meglio i cittadini), dato che prima della Guerra di Indipendenza si era formata una comunità di agricoltori, senza stratificazioni sociali che, schiavi esclusi, godevano dei pieni diritti civili e politici ed erano proprietari delle terre che coltivavano, legando la formazione della ricchezza alla proprietà fondiaria<sup>20</sup>. Infatti, nelle colonie britanniche d'America esisteva una comunità di agricoltori, bianchi e non cattolici, che avevano ottenuto la cittadinanza britannica (*Naturalization Law*) durante l'epoca coloniale<sup>21</sup>. I coloni inurbati, invece, si erano arricchiti con lo scambio mercantile, rifuggendo però di sottostare alle gravose imposte del Parlamento britannico, ma legando il valore della cittadinanza e della libertà all'idea della rappresentanza, in particolare per quel che concerneva l'imposizione fiscale. Essi credevano che sottostare a imposte decise dal Parlamento li avrebbe resi "schiavi tributari"<sup>22</sup>. L'idea di "riservare" i diritti di libertà attribuendoli ai soli maschi liberi di religione protestante rappresenta uno dei punti rimasti di dubbia comprensibilità ancora oggi, a meno che non si consideri il costante e profondo elemento culturale che lega i coloni ribelli ai moderni americani nella considerazione che il fondamento della libertà risieda nella garanzia della proprietà privata<sup>23</sup>.

Il tema del retaggio della schiavitù è rilevante per delineare la questione, soprattutto in relazione alla difficoltà dei cittadini afroamericani di liberarsi della disuguaglianza di partenza legata alle loro origini. In un certo senso è ragionevole affermare, con il supporto di fatti evidenti, che lo sfruttamento lavorativo si ripete anche ora, grazie all'applicazione di strumenti decisori automatizzati, ai quali è delegata l'organizzazione delle mansioni lavorative nella gestione industriale moderna, soprattutto in ambiti gestionali complessi e spersonalizzati.

Il mito della Rivoluzione francese, invece, pretese di contrapporre una sorta di fantasioso Ancien Régime<sup>24</sup>, che giustificasse le scelte politiche (e giuridiche) successive. Si trattò di una raffigurazione manichea e mistificante per far breccia sul passato, ricostruirlo e sostituirlo forzatamente con il nuovo corso<sup>25</sup>, spiccatamente

---

<sup>20</sup> P. Greco, *La Scienza e l'Europa, Dal Seicento all'Ottocento*, L'Asino d'Oro Edizioni, Roma, 2016.

<sup>21</sup> P. Greco, *La Scienza e l'Europa*, cit., p. 304.

<sup>22</sup> A. Taylor, *Rivoluzioni Americane. Una storia continentale 1750 -1804*, trad. it. D. Di Nunno, Einaudi, Torino, 2017, p. 121 ss. Specifica l'autore: "Nel maggio 1764, i membri dell'assemblea di Boston ribadirono: "Se vengono applicate su di noi tasse di qualsiasi tipo senza mai aver avuto un rappresentante legittimo laddove sono state disposte, non vediamo ridotti dall'attributo di sudditi liberi alla miserabile condizione di schiavi tributari?"

<sup>23</sup> E. Foner, *Storia della libertà americana*, trad. it. A. Merlino, Donzelli Editore, Roma, 2010, p. 51 ss.

<sup>24</sup> F. Furet, voce "Ancien Régime", in *Dizionario critico della Rivoluzione francese*, a cura di F. Furet e M. Ozouf, Bompiani, Milano, 1988, p. 559-569.

<sup>25</sup> H. Brustin, *Rivoluzionari, Antropologia politica della Rivoluzione francese*, Laterza, Roma-Bari, 2016, p. 13 ss.

politico invece che economico ovvero tecnologico. Tuttavia la questione sulla libertà personale e uguaglianza non poteva rimanere inascoltata.

## 2. Rivoluzioni dei diritti e rivoluzione industriale: l'avvento dell'automazione

L'esperienza della Rivoluzione industriale ha dimostrato che presupposto delle rivendicazioni di uguaglianza e diritti è da ricondurre all'introduzione dell'automazione nel processo produttivo che ha favorito l'avvento della manifattura dei beni industriali con la creazione delle fabbriche e lo spostamento della produzione dalle rive dei fiumi e vicino alle fonti d'acqua ai centri urbani. Tale migrazione produttiva ha creato le condizioni per la nascita delle città industriali, grazie alla invenzione della macchina a vapore di James Watt<sup>26</sup>. Come è noto, tale innovazione comportò la trasformazione della società a causa dello spostamento di masse di persone dalle campagne alle città, ma poste in una realtà del tutto straniente, le fabbriche, cancellando mentalità, comportamenti e consuetudini legate alla precedente vita agricola e in questa radicate nella cultura e memoria collettiva<sup>27</sup>.

Da un punto di vista del rapporto del diritto con la tecnologia, tale cambiamento ha avuto un forte impatto giuridico e sociale in quanto l'evoluzione tecnologica promuove una spinta unificante attraverso l'uniformazione delle regole di condotta del nuovo processo produttivo. Infatti, siffatta transizione rappresenta un passaggio verso un sistema articolato di norme, le quali, in funzione di precetti tecnici, impongono a ciascun soggetto coinvolto di omologare il proprio comportamento in funzione di standard tecnici che assumono valore legale<sup>28</sup>.

Questo contesto ci consente di effettuare una comparazione storicamente diacronica tra l'automazione introdotta nel XIX secolo nelle fabbriche inglesi e l'automazione algoritmica attuale e tale comparazione riguarda la misurazione del tempo durante lo svolgimento della mansione lavorativa. Come già accaduto in passato, il primo ambito ove si manifesta e sviluppa l'innovazione dell'attività umana è l'ambito lavorativo e questa circostanza è valida anzitutto per quel che concerne l'automazione: da un lato l'espletamento di un lavoro, castigo biblico, consente agli individui di vivere una vita indipendente, dall'altro lato automatizzare il lavoro consente agli imprenditori e ai produttori una serie di risparmi a discapito dei lavoratori, il

---

<sup>26</sup> P. Greco, *op. cit.*, p. 321.

<sup>27</sup> P. Mantoux, *The industrial revolution in the eighteenth century: An outline of the beginnings of the modern factory system in England*, Routledge, Abingdon-on-Thames, 2013, *passim*; D.S. Landes, *The unbound Prometheus: technological change and industrial development in Western Europe from 1750 to the present*, Cambridge University Press, Cambridge, 2003, *passim*.

<sup>28</sup> Il caso cui ci si riferisce riguarda l'implementazione di norme relative alla circolazione stradale, conseguenti all'acquisto di massa di automobili, tuttavia tale fenomeno avviene ogni volta che una innovazione tecnologica viene utilizzata massivamente da utenti e/o consumatori (G. Calabresi, E. Al Mureden, *Driveless Cars*, Il Mulino, Bologna, 2021, p. 27).

cui ruolo viene ricondotto a quello di “mera macchina”<sup>29</sup>, realizzando così la maledizione marxista del “lavoro alienato”<sup>30</sup>.

La misurazione del tempo del lavoro, in prospettiva di risparmi sui costi del lavoro stesso, è una delle preoccupazioni più importanti della gestione di un'organizzazione complessa come quella aziendale: l'automazione del procedimento produttivo del XIX secolo slega il lavoro dalla durata della luce diurna (grazie all'elettricità) e alle avversità meteorologiche, come invece continuava ad accadere in agricoltura, mentre l'automazione algoritmica estremizza questa situazione: l'algoritmo può gestire i processi produttivi senza controllo umano, realizzando quel tipo di alienazione evocata da Marx del singolo lavoratore estraniato dagli altri lavoratori<sup>31</sup>.

Applicando la teoria marxista sulla durata delle giornate lavorative<sup>32</sup>, *il tempo* come ora, è possibile osservare un effetto paradossale: da un lato il capitalista aumenta i suoi profitti forzando la capacità lavorativa dei lavoratori e aumentando il plusvalore della produzione, dall'altro lato tale metodo si traduce in rendimenti marginali, poiché l'allungamento dei turni esaurisce i lavoratori. Ciò è vero sia per quel che concerne il capitalista manifatturiero sia per la piattaforma digitale che impiega i lavoratori nello svolgimento delle mansioni gestite dagli algoritmi<sup>33</sup>. In entrambe le situazioni, l'aumento del plusvalore prodotto dalla forza lavoro impegnata comporta il controllo fisico dei lavoratori. Se in passato detto controllo era reso problematico in quanto, a differenza di altri fattori produttivi coinvolti nella produzione, il potere del lavoro è (*rectius, era*) costituito dalla sua incarnazione nelle persone: i lavoratori, che a loro volta mantengono (*rectius, mantenevano*) il loro potere di resistere nell'essere trattati alla stregua di merci<sup>34</sup>. Nei tempi della “Rivoluzione digitale” siffatta corrispondenza tra fisicità e lavoro è messa fortemente in crisi dalla solitudine dello svolgimento della mansione lavorativa: infatti il lavoratore avente quale riferimento la piattaforma gestita da algoritmi si trova ad essere isolato dagli altri colleghi e ha quale uno mezzo di relazione diretta l'applicazione informatica che consente l'interazione tra il medesimo e l'algoritmo.

Quanto descritto produce un ulteriore effetto paradossale, il quale riguarda la circostanza relativa allo sfruttamento gratuito del lavoro degli schiavi, soprattutto per quel che concerne gli Stati Uniti<sup>35</sup>, messo in crisi dall'automazione meccanica

---

<sup>29</sup> K. Marx, *I manoscritti economico-filosofici del 1844*, a cura di N. Bobbio, Einaudi, Torino, 2004, *passim*.

<sup>30</sup> K. Marx, *I manoscritti economico-filosofici*, cit.; G. Claeys, *Marx e il marxismo*, trad. it. A. Manna, Einaudi, Torino, 2018, p. 58 ss.

<sup>31</sup> K. Marx, *I manoscritti economico-filosofici*, cit.; G. Claeys, *Marx e il marxismo*, cit.

<sup>32</sup> K. Marx, *il Capitale*, Editori Riuniti, Roma, 1989, cap. 10.

<sup>33</sup> *Supra*, Cap. IV, par. 5.

<sup>34</sup> J. Woodcock, *The algorithmic panopticon at Deliveroo: Measurement, precarity, and the illusion of control*, in *Ephemera* 20.3, (2020), pp. 67-95.

<sup>35</sup> J.R. Hummel, *Emancipating slaves, enslaving free men: a history of the American Civil War*, Open, Chicago, 2014, p. 37 ss.

e lo sviluppo delle fabbriche. Tale tema torna d'attualità attraverso la discriminazione algoritmica effettuata nel mondo del lavoro contemporaneo dove i lavoratori vengono privati di adeguato riconoscimento giuridico e salariale. La discriminazione algoritmica trova il suo fondamento a partire dal mondo del lavoro perché, come dimostra l'esperienza statunitense, essa consente il perpetuarsi di distinzioni categoriali e giuridiche che riflettono quelle collegate al lavoro<sup>36</sup>.

La Rivoluzione industriale muta il paradigma dell'uguaglianza e, a specchio, anche quello della discriminazione in quanto l'avvento del capitalismo aveva provocato un cambiamento reale nella società e negli stili di vita<sup>37</sup>. Nel discorso filosofico l'uguaglianza acquisiva una centralità nuova che si staccava dalla prospettiva religiosa e filosofica e assumeva quella sociale e politica incarnata nelle grandi Rivoluzioni del tempo, cioè quella americana e francese<sup>38</sup>.

Uguaglianza, disuguaglianza e, quindi, discriminazione uscivano dalla condizione formale della mera schiavitù, la quale tuttavia non era ancora affatto debellata, ed entravano nel novero dei fatti sociali sui quali iniziò a formarsi un'opinione pubblica più ampia rispetto al vecchio ordine aristocratico e religioso. Si iniziò a riflettere che la discriminazione era legata alla povertà, all'applicazione delle tecnologie che frantumavano le categorie sociali, indebolendole.

Frammentazione e indebolimento di siffatte categorie sembrerebbero rinnovarsi anche nei tempi odierni, in cui la tecnologia potenzia la discriminazione verso le fasce più vulnerabili nei confronti delle quali essa non sembra più essere sradicabile, ma al contrario appare rafforzarsi perché consente l'emersione di "elementi espansivi di oggettiva impersonale eguaglianza, rispetto a ogni tipo di differenza individuale o di genere"<sup>39</sup>.

Si avrebbe quindi un effetto irragionevole secondo cui l'uguaglianza si tramuterebbe in una forma di impersonalità in quanto, per raggiungerla, occorrerebbe ignorare, o addirittura annullare, le caratteristiche individuali che rendono irripetibile ciascuna persona<sup>40</sup>.

Tuttavia, è in questo contesto di serializzazione omogenea delle individualità che si verificano fattispecie discriminatorie, in particolare nei trattamenti automatizzati dei dati personali. In questo senso gli algoritmi decisori automatizzati trovano la giustificazione dell'efficienza nel loro uso proprio nella serialità delle questioni e sovrapposibilità delle fattispecie. Infatti, nella prospettiva della automazione algoritmica si prospetta una sorta di indifferenza, ovvero una pretesa neutralità, sulla possibile origine della disuguaglianza, sia fisica ovvero morale, perché la nomencla-

---

<sup>36</sup> C. Fuchs, *Digital Labour and Karl Marx*, Routledge, Abingdon-on-Thames, 2014. *Supra*, Cap. IV, par. 5.

<sup>37</sup> A. Schiavone, *Eguaglianza*, Einaudi, Torino, 2019, p. 117.

<sup>38</sup> A. Schiavone, cit., p. 118.

<sup>39</sup> A. Schiavone, cit., p. 292 ss.

<sup>40</sup> M. Ainis, *op. cit.*

tura delle scelte elaborate a monte dall'autore dell'algoritmo riguardano una catalogazione, e quindi una classificazione basata sull'elaborazione dei concetti. Tale procedimento consente di scegliere nel corso della formazione dell'algoritmo come classificare ciascun elemento partecipe del procedimento sulla base degli scopi cui l'algoritmo è diretto ovvero delle convinzioni del programmatore. Siffatta operazione non può essere neutrale, ma segue dei criteri di catalogazione secondo lo scopo che si intende ottenere. Dal momento che l'operazione in parola non è neutrale, né può esserlo, essendo basata su scelte, essa contiene in nuce potenziali discriminazioni. Su tale questione certa dottrina<sup>41</sup> ha indagato se e come sia possibile adottare misure di elementi differenti relative alla *fairness*<sup>42</sup>, esaminando in particolare le relazioni matematiche tra di essi. Per esempio, tra la calibrazione dei gruppi, il bilanciamento delle classi positive ed il bilanciamento delle classi negative dimostrando che "eccetto in ristretti casi specifici" non vi siano condizioni che possano soddisfare simultaneamente le condizioni trattate nell'esperimento<sup>43</sup>. Altri studi hanno affrontato le questioni relative sulle incompatibilità tra i criteri di *fairness*, i quali in certi contesti (come i test psicologici applicati alle recidive) possono portare a una significativa discriminazione indiretta quando la recidiva stessa si differenzia tra i gruppi esaminati<sup>44</sup>. Altresì, nell'analisi del rischio criminale è stato osservato che è generalmente "impossibile massimizzare l'accuratezza e la *fairness* allo stesso tempo, ed impossibile soddisfare contemporaneamente tutti i tipi di *fairness*"<sup>45</sup>.

Ci si potrebbe domandare se in condizioni di trasformazione tecnologica, come avvenuto a metà del XVIII secolo e come sta avvenendo ora, si possa intravedere l'origine della discriminazione in considerazione della vulnerabilità dei soggetti coinvolti nell'utilizzo dell'innovazione apportata dall'automazione, tanto nella prima Rivoluzione industriale, quanto in quella contemporanea legata all'utilizzo dell'intelligenza artificiale. In entrambi i casi vi è una subalternità delle classi vulnerabili che non possono, né sono (*erano*) in grado di sottrarsi a tale mutamento di paradigma tecnologico. Si tratta di una condizione dovuta all'appartenenza alle categorie vulnerabili e, al contempo, è la stessa vulnerabilità che costringe tali parti deboli ad essere subalterne.

Nei Paesi afferenti alla Western Legal Tradition, in particolare negli Stati Uniti,

---

<sup>41</sup> P. Garg, J. Villasenor, V. Foggo, *Fairness Metrics. A Comparative Analysis*, in 2020 *IEEE International Conference on Big Data (Big Data)* (pp. 3662-3666).

<sup>42</sup> Termine che si è deciso di lasciare nella lingua originale degli studi citati in quanto contrapposto al "bias" (si veda *supra*, Cap. IV, par. 1). Esso potrebbe essere altresì tradotto come "correttezza" ovvero "equità", indicativi di una contrapposizione alla discriminazione.

<sup>43</sup> J. Kleinberg, S. Mullainathan, M. Raghavan, *Inherent trade-offs in the fair determination of risk scores*, 2016, arXiv preprint arXiv:1609.05807.

<sup>44</sup> A. Chouldechova, *Fair prediction with disparate impact: A study of bias in recidivism prediction instruments*, 2017, <https://arxiv.org/abs/1703.00056>.

<sup>45</sup> R. Berk, H. Heidari, S. Jabbari, M. Kearns, A. Roth, *Fairness in criminal justice risk assessments: The state of the art*, in *Sociological Methods & Research*, 2021, p. 3-44.