

GenIUS

RIVISTA DI STUDI GIURIDICI
SULL'ORIENTAMENTO SESSUALE E L'IDENTITÀ DI GENERE

NICOLA LETTIERI

La discriminazione nell'era
delle macchine intelligenti.
Modelli possibili di analisi, critica e tutela

PUBBLICAZIONE TELEMATICA SEMESTRALE REGISTRATA PRESSO IL TRIBUNALE DI BOLOGNA · ISSN 2384-9495

online first

24 ottobre 2022

La discriminazione nell'era delle macchine intelligenti. Modelli possibili di analisi, critica e tutela

Sommario

1. Introduzione. - 2. Macchine intelligenti e discriminazione. Anatomia di un fenomeno complesso. - 3. Critica e analisi della discriminazione algoritmica: due prospettive di ricerca. - 4. *Computational law*: gli algoritmi da problema a soluzione. - 5. Algoritmi contro la discriminazione algoritmica: una sperimentazione. - 6. Conclusioni.

Abstract

L'articolo discute le nuove forme di discriminazione oggi generate dall'ingresso di sistemi algoritmici all'interno di processi decisionali che incidono su interessi giuridicamente rilevanti. Fulcro dell'analisi è una riflessione critica intorno al modo in cui il mondo del diritto concettualizza il fenomeno e progetta strategie regolative per contrastarlo. Con riguardo al primo dei due aspetti, l'idea è che sia necessario ripensare la discriminazione mediata dalle macchine attingendo alle prospettive, ai metodi e alle evidenze empiriche provenienti dai settori dei *Critical data studies* e delle Scienze sociali computazionali. La seconda parte del lavoro muove dalla considerazione della scarsa efficacia che spesso caratterizza l'operato dei rimedi giuridici tradizionali rispetto a fenomeni che, come la discriminazione algoritmica, sono fortemente connotati in termini tecnologici. Il punto di approdo è la tecnoregolazione, paradigma normativo presentato non solo in termini astratti, ma anche attraverso il riferimento a un progetto di ricerca che attinge ai metodi delle scienze sociali computazionali per esplorare approcci innovativi allo studio e al contrasto delle forme di discriminazione che hanno luogo nella *gig economy*.

The article discusses the new forms of discrimination produced by the spread of algorithmic systems into decision-making processes impacting legally relevant interests. The work revolves around a critical reflection on how the legal world conceptualizes the phenomenon and designs regulatory strategies to cope with it. As to the first aspect, we dwell on the need to rethink machine-mediated discrimination by drawing on the perspectives, methods and empirical evidence from Critical data studies and computational social sciences. The second part of the work cues from the lack of effectiveness that often characterizes traditional legal remedies in facing phenomena that, like algorithmic discrimination, are strongly connoted in technological terms. The arrival point is techno regulation, a normative

* Docente a contratto, Università degli studi del Sannio. Relazione al Convegno "L'algoritmo alla prova del caso concreto: stereotipi, serializzazione, discriminazione" ospitato dall'Università di Ferrara in data 6.4.2022. Contributo non sottoposto a referaggio a doppio cieco.

paradigm also discussed by referring to an ongoing research project that draws on computational social sciences methods and explores innovative approaches to study and fight algorithm-driven discrimination in the gig economy.

1. Introduzione

Quello della discriminazione mediata dagli algoritmi è un tema che occupa senza dubbio una posizione di primo piano tra le tante questioni sollevate dall'evoluzione in senso tecnologico della nostra società¹. Complice una serie di casi di cronaca ampiamente ripresi dai media e dibattuti, gli ultimi anni sono stati segnati dalla presa di coscienza che l'impiego di sistemi algoritmici² per assumere decisioni in grado di incidere sulla sfera giuridica individuale porta con sé rischi considerevoli di lesione per diritti di fondamentale rilevanza come il diritto alla non-discriminazione.

L'attenzione riservata al fenomeno, a onor del vero, è già stata considerevole. La discriminazione operata *da* (o *attraverso*) le macchine ha già alimentato una quantità notevole di analisi e riflessioni in ambiti diversi della scienza giuridica dagli studi giuslavoristici al diritto costituzionale e penale passando per la filosofia del diritto³. A tutto ciò si aggiunge un numero ormai consistente di interventi normativi adottati a livelli ordinamentali diversi per contrastare l'insorgere di dinamiche discriminatorie connesse all'uso di sistemi di *algorithmic decision making*.

Al netto dei risultati conseguiti sinora, tuttavia, la lotta alla discriminazione trova ancora sul suo percorso ostacoli significativi in vario modo connessi alla complessità che caratterizza il fenomeno in termini fattuali prima ancora che giuridici.

La prima difficoltà con cui è necessario confrontarsi è di carattere conoscitivo. I meccanismi, le forme e le implicazioni dell'ingiustizia algoritmica sono difficili da identificare e comprendere nel loro concreto svolgersi e questo, come vedremo, per ragioni che vanno anche al di là della connotazione tecnologica del fenomeno. Un secondo ordine di problemi, poi, si colloca sul piano ordinamentale dove i rimedi giuridici tradizionali fanno fatica da soli a rispondere con la necessaria efficacia a processi rapidissimi, implementati da infrastrutture tecnologiche inaccessibili espressione di quella "planetary scale computation"⁴ che pone oggi dinanzi ai nostri occhi modelli del tutto nuovi di geografia istituzionale, politica ed economica.

-
- 1 Sul tema, a titolo esemplificativo, si veda C. O'Neil, *Weapons of math destruction: How big data increases inequality and threatens democracy*, Portland, Broadway Books, 2016.
 - 2 Una definizione operativa di *algorithmic decision making*, adottabile anche qui a scopi operative, è quella offerta nello studio *Understanding algorithmic decision-making: Opportunities and challenges*, predisposto nel 2019 dalla Scientific Foresight Unit del Panel scientifico del Parlamento europeo: "While algorithms are hardly a recent invention, they are nevertheless increasingly involved in systems used to support decision-making. These systems, known as 'ADS' (*algorithmic decision systems*), often rely on the analysis of large amounts of personal data to infer correlations or, more generally, to derive information deemed useful to make decisions".
 - 3 La letteratura da prendere in considerazione è già molto ampia. Una per ogni aspetto M. Barbera, *Discriminazioni algoritmiche e forme di discriminazione*, in *Labour & Law Issues*, 2021, 7 (1), pp. I-1.; P. Dunn, *Moderazione automatizzata e discriminazione algoritmica: il caso dell'hate speech*, in *Rivista italiana di informatica e diritto*, 2022, 4 (1), p. 12; V. Faggiani, *La lotta contro il crimine attraverso gli algoritmi: contraddizioni e profili di (in)costituzionalità dell'applicazione dell'IA alla giustizia penale*, in *Respuesta institucional y normativa al crimen organizado. Perfiles estratégicos para una lucha eficaz*, a cargo de F. Garrido Carrillo, Navarra, 2021, pp. 245-282; A. Simoncini, S. Suweis, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto*, 2019, 8 (1), pp. 87-106.
 - 4 B.H. Bratton, *The stack: On software and sovereignty*, Cambridge, MIT Press, 2016.

In uno scenario di questo tipo, il mondo del diritto si trova di fronte a sfide che trascendono la pur imprescindibile attività di inquadramento giuridico - *de jure condito* e *de jure condendo* - delle nuove forme di discriminazione. La particolare natura del fenomeno spinge in effetti anche in altre direzioni: se, per un verso, appaiono necessarie forme più raffinate di analisi empirica delle dinamiche discriminatorie oggetto di regolazione, per altro verso è ormai evidente l'esigenza di esplorare nuove soluzioni regolative al di fuori dei paradigmi tradizionali.

Il contributo interviene su questi temi articolando una riflessione che trae spunto da una lunga attività di ricerca e sperimentazione condotta ai confini tra diritto e *computer science*⁵. Dopo un breve inquadramento della discriminazione algoritmica e delle sue diverse manifestazioni concrete, ci si concentrerà sulla prima delle sfide evocate proponendo l'apertura di un nuovo fronte di ricerca alimentato dal dialogo con due aree di studio - le scienze sociali computazionali e i *Critical data studies* - che sembrano poter contribuire in maniera significativa alla comprensione delle meccaniche di fondo della discriminazione e alla loro valutazione critica.

L'attenzione si sposterà quindi sulle questioni regolative esaminando l'idea di utilizzare ICT, dati e algoritmi per offrire risposte più efficaci ai problemi di controllo e governo giuridico delle interazioni sociali, personali ed economiche mediate dalla tecnologia. Si farà riferimento in questa prospettiva a una sperimentazione che attinge ai metodi delle scienze sociali computazionali per esplorare approcci innovativi allo studio e al superamento delle forme di discriminazione che hanno luogo nello scenario della *gig economy*. L'articolo si conclude con alcune considerazioni di carattere generale sui cambiamenti che l'evoluzione in senso algoritmico della società sembra suggerire al giurista sul piano culturale, scientifico e metodologico.

2. Macchine intelligenti e discriminazione. Anatomia di un fenomeno complesso

Nel febbraio 2020, con una pronuncia destinata senza dubbio ad alimentare il dibattito sull'uso delle nuove tecnologie, la Corte dell'Aja ha dichiarato l'illegittimità di SyRI, algoritmo sviluppato su impulso del ministero degli affari sociali olandese per combattere le frodi al sistema previdenziale nazionale. Alimentato da archivi pubblici contenenti grandi quantità di dati personali e sensibili, SyRI ha per diversi anni valutato la propensione dei cittadini a frodare lo Stato assegnando a ciascun individuo (spesso del tutto estraneo a intenti frodati) un indice di rischio usato dalle autorità per disporre controlli fiscali⁶.

5 L'attività nasce da una collaborazione con il Dipartimento di Informatica dell'Università degli Studi di Salerno e attinge alle riflessioni maturate nell'ambito di due diversi corsi universitari: il corso "Intelligenza computazionale e tecnoregolazione: paradigmi, metodi, strumenti" tenuto insieme alla prof.ssa Delfina Malandrino e al prof. Rocco Zaccagnino nell'ambito del dottorato di ricerca in Informatica, e il corso "Diritto e scienze sociali computazionali" tenuto presso il Dipartimento di Diritto, Economia, Management e Metodi Quantitativi dell'Università degli Studi del Sannio. L'attività ha condotto allo sviluppo di GigAdvisor, prototipo di una piattaforma accessibile via web (<http://gigadvisor.herokuapp.com/GigAdvisor/home/>) e app mobile concepita come "laboratorio virtuale" per esperimenti collegati tanto al paradigma dei *Critical Data Studies*, quanto a quello della tecnoregolazione.

6 A portare il governo olandese davanti al giudice è stata una coalizione di associazioni, fra cui il comitato di giuristi per i diritti umani dei Paesi Bassi Juristen voor Mensenrechten, la fondazione Privacy First, un sindacato e un'associazione di consumatori. L'uso di SyRI, pure avallato dal Parlamento olandese, ha suscitato critiche sin da quando il governo ha diffuso la notizia che l'algoritmo sarebbe stato utilizzato per passare in rassegna i quartieri poveri di Rotterdam, Eindhoven e

Oltre a rilevare l'impossibilità di stabilire con certezza natura e consistenza dei dati analizzati, la sentenza ha stigmatizzato l'opacità dei criteri impiegati dall'algoritmo per valutare il rischio di frode. In queste circostanze la Corte ha ravvisato la compressione di due garanzie fondamentali offerte dalla Convenzione Europea dei Diritti Umani consistenti, rispettivamente, nel riconoscimento del diritto alla vita privata e nel divieto di discriminazione. Letta in profondità, la vicenda dell'algoritmo *SyRI* rappresenta qualcosa di più di un precedente giurisprudenziale ma l'indice di un processo molto più esteso che coinvolge la società dell'informazione in termini ampi e generali.

Nel mondo del lavoro, per citare uno dei contesti più dibattuti⁷, l'utilizzo di sistemi di decisione algoritmica per compiere scelte gestionali di notevole rilevanza come quelle in materia di turni, retribuzione, misure disciplinari, valutazione delle *performance*⁸, ha già mostrato di poter alimentare rischi significativi di discriminazione tra i lavoratori. Le tecniche di analisi dei dati utilizzate per classificare i dipendenti, formulare previsioni e assumere decisioni⁹ mostrano una sensibilità tale alle caratteristiche intrinseche delle informazioni elaborate da rendere molto elevato il rischio di aberrazioni ed errori anche involontari¹⁰.

Fenomeni simili si materializzano in ambito giudiziario dove l'introduzione di ausili computazionali che sfruttano il calcolo predittivo ha dato prova di potersi tradurre in lesioni del diritto alla difesa o di quello a un giusto processo. La vicenda ormai nota di COMPAS¹¹, piattaforma pensata per supportare i magistrati nella valutazione dei rischi di recidiva, rappresenta da questo punto di vista solo l'esempio più noto delle questioni sollevate dall'impiego di modelli previsionali nell'amministrazione della giustizia.

Quanto detto lascia intuire la complessità del rapporto tra algoritmi e dati e, di conseguenza, le

Haarlem sottoponendo alla profilazione anche cittadini che non percepivano sussidi purché fossero residenti nelle zone atenzionate.

- 7 Non è un caso che la Proposta di regolamento del Parlamento europeo e del Consiglio recante norme armonizzate sull'intelligenza artificiale (COM(2021) 206 final) presentata nell'aprile 2021 identifichi nei sistemi intelligenti utilizzati nel mondo del lavoro, una categoria di sistemi "ad alto rischio" capaci, in quanto tali, di porre "rischi significativi per la salute e la sicurezza o per i diritti fondamentali delle persone". Il punto 36 della Proposta elenca tra i sistemi "ad alto rischio" i "sistemi di IA utilizzati nel settore dell'occupazione, nella gestione dei lavoratori e nell'accesso al lavoro autonomo".
- 8 Frequenti sono i casi di discriminazione basata su ragioni di razza, sesso, religione o posizione geografica. Per un'analisi caratterizzata da un interessante approccio interdisciplinare si v., tra gli altri, P.T. Kim, *Data-driven discrimination at work*, in *William & Mary Law Review*, 2016, 58, pp. 857-936.
- 9 Un ruolo centrale è giocato in questo scenario dal *machine learning*, ambito dell'intelligenza artificiale, nel quale converge un insieme di metodi di apprendimento automatico pensati per estrarre regolarità e relazioni all'interno di grandi quantità di dati. L'output è costituito da classificazioni, generalizzazioni e previsioni che possono tanto rappresentare un risultato in sé, quanto costituire il punto di partenza per azioni e decisioni più complesse. Per una panoramica delle prospettive dischiuse dal *machine learning* si v. P. Domingos, *L'Algoritmo Definitivo: La macchina che impara da sola e il futuro del nostro mondo*, Torino, Bollati Boringhieri, 2016.
- 10 Per quanto ampi, i *dataset* offrono rappresentazioni limitate del mondo rispetto alle quali è essenziale chiedersi: cosa viene quantificato, archiviato e ordinato? Cosa viene scartato? La produzione, raccolta e interpretazione dei dati, d'altra parte, non sono solo attività tecnicamente complesse, ma anche il frutto di pratiche sociali contingenti che sottendono visioni del mondo e codificano all'interno di scelte tecniche (formato e contenuto dei dati, campi e metriche da usare per le misurazioni) assunti e pregiudizi, obiettivi conoscitivi e applicativi. Per una riflessione critica sul tema dell'obiettività dei dati, si veda L. Gitelman, (ed.), *Raw data is an oxymoron*, Cambridge, MIT Press, 2013.
- 11 La vicenda giudiziaria originata dall'utilizzo della piattaforma COMPAS (acronimo di Correctional Offender Management Profiling for Alternative Sanctions) per la valutazione del rischio di recidiva è ampiamente dibattuta. Per un approfondimento sul punto si veda, con particolare riferimento alle implicazioni per il diritto costituzionale, A. Simoncini, S. Suweis, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, cit.

tante criticità che possono affliggere i meccanismi di decisione e regolazione algoritmica. I rischi di una compressione, anche involontaria, di diritti fondamentali possono derivare da una moltitudine di fattori.

Le istituzioni europee sembrano aver acquisito sempre maggiore consapevolezza delle insidie che si celano dietro i sistemi di decisione algoritmica, individuando soluzioni¹² da accogliere senza dubbio con favore: l'affermazione della necessità dello *human oversight*, del controllo umano sui sistemi di decisione algoritmica o, ancora, la proiezione verso il modello della *explainable AI*¹³, di una Intelligenza artificiale di cui gli esseri umani possono comprendere le decisioni o le previsioni, sono tutti passi avanti importanti. Il punto è che la discriminazione mediata dagli algoritmi è, in termini fattuali, un fenomeno più complesso di quanto queste scelte sembrino supporre.

A ben vedere, i rischi indotti dai sistemi di *algorithmic decision-making* vanno oltre l'ipotesi - ancora apparentemente centrale nella visione del legislatore europeo - di una discriminazione operata da una *singola* decisione automatica in danno di una *singola* persona. Essi includono anche altri scenari in cui le dinamiche di tipo discriminatorio favorite dalle macchine investono *più soggetti* spostandosi dal piano squisitamente individuale a quello collettivo e sociale.

Gli algoritmi per la selezione dei contenuti da mostrare agli utenti dei *social media*, per fare un esempio, svolgono un ruolo determinante nella diffusione di modelli culturali e stereotipi che incidono in maniera significativa sul materializzarsi di fattispecie di tipo discriminatorio - si pensi all'istigazione all'odio razziale o etnico¹⁴ - che gli ordinamenti nazionali e internazionali provano da tempo a contrastare. In casi di questo tipo, inoltre, il contributo degli algoritmi al prodursi di nuove forme di discriminazione si somma ad altri fattori ugualmente rilevanti sul piano causale che vanno dai *bias* cognitivi che condizionano il *decision making* alla topologia delle reti sociali in cui gli individui sono immersi per arrivare alla meccanica delle interazioni collettive.

Lo scenario descritto trova conferma in una pluralità di evidenze scientifiche. La rilevanza del modo in cui le scelte individuali interagiscono tra di loro è ben nota sin dal pionieristico studio sulla segregazione residenziale¹⁵ pubblicato dal premio Nobel per l'economia Thomas Schelling negli anni

-
- 12 Nel corso degli ultimi anni, il legislatore europeo ha riservato a questi temi un'attenzione crescente attraverso una serie di provvedimenti in vario modo collegati all'IA e ai sistemi di decisione algoritmica nei quali il rischio di discriminazione mediata dalle macchine occupa una posizione centrale. Riferimenti significativi in questa direzione si v., Commissione Europea, *On Artificial Intelligence. A European approach to excellence and trust, White Paper*, (COM (2020)) 65; *Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence* (COM (2021) 206 final); *Proposal for a directive of the European Parliament and of the Council on improving working conditions in platform work* (COM (2021) 762 final; Fundamental Rights Agency, *Preparare un giusto futuro — L'intelligenza artificiale e i diritti fondamentali*.
- 13 Il "right to explanation" si collega ad un tema di ricerca estremamente ampio, attuale e gravido di implicazioni, quello della "explainable AI" e cioè a quell'insieme di metodi e tecniche orientate a far sì che i risultati prodotti dai sistemi di intelligenza artificiale possano essere compresi dagli umani. Sul punto si veda, tra gli altri, D. Gunning, D. Aha, *DARPA's explainable artificial intelligence (XAI) program*, in *AI magazine*, 2019, 40 (2), pp. 44-58; B. Goodman, S. Flaxman, *European Union regulations on algorithmic decision-making and a "right to explanation"*, in *AI magazine*, 2017, 38 (3), pp. 50-57.
- 14 Il riferimento è alle fattispecie contemplate nell'art. 3, co. 1 lett. a) e b) della l. 654/1975: *propaganda idee fondate sulla superiorità o sull'odio razziale o etnico; istigazione alla commissione di atti di discriminazione per motivi razziali, etnici, nazionali o religiosi; incitamento alla commissione di violenza o di atti di provocazione alla violenza per motivi razziali, etnici, nazionali o religiosi*.
- 15 Il modello riesce ad offrire una visione penetrante del fenomeno a partire da una simulazione strutturalmente molto semplice, destinata ad evolvere sulla base delle scelte compiute da agenti artificiali che operano in funzione di un criterio di decisione tollerante (la "regola della felicità"): ciascuno di essi si sposta dal proprio luogo di residenza solo se la percentuale di vicini di colore diverso dal proprio supera il 70% all'interno del vicinato (rappresentato simbolicamente dalle otto celle circostanti). Nel modello, il formarsi di gruppi compatti e isolati tra loro di individui della stessa razza è il risultato

Settanta. Utilizzando un'innovativa tecnica di simulazione computazionale di dinamiche sociali basata sui automi cellulari¹⁶, Schelling mostra come fenomeni di tipo discriminatorio possano emergere in maniera spontanea e del tutto controintuitiva dalla dinamica dell'interazione che si sviluppa tra scelte individuali non qualificabili come razziste.

Oggi, a distanza di anni, uno scenario analogo emerge dallo studio dei *social media*. Un numero ormai consistente di ricerche¹⁷ ha confermato la rilevanza della struttura di interazione sociale¹⁸ rispetto alla genesi di fenomeni discriminatori offrendo prove significative in merito a un ruolo degli algoritmi (gli algoritmi di *ranking* delle news, gli account automatici che fingono di essere persone vere e pubblicano post sui *social network*, gli algoritmi alla base dei motori di ricerca) che non è quello di operare in maniera diretta discriminazioni, quanto quello di catalizzare i processi sociali e culturali che conducono a quest'ultima.

Oltre ad avere rilievo sul piano scientifico, quanto detto merita attenzione anche dal punto di vista del diritto che fatica a elaborare risposte ordinamentali adeguate alla complessità dei processi appena descritti, una complessità che è difficile da comprendere prima ancora che da regolare. Tutto questo spinge verso il dialogo con nuove forme di analisi critica e ricerca empirica che possano contribuire a radicare l'azione regolativa in una comprensione più profonda dei meccanismi fattuali che condizionano l'emergere e l'evolvere dei fenomeni di discriminazione mediata dalle macchine.

3. Critica e analisi della discriminazione algoritmica: due prospettive di ricerca

Quanto descritto nei paragrafi precedenti ha offerto un'idea di come la comprensione del *modus operandi* dei sistemi di decisione algoritmica sia una condizione fondamentale per valutare confini, rischi

emergente dalle interazioni tra gli agenti e dai vincoli fissati per esse. La segregazione, per intendersi, è una proprietà macro del sistema sociale simulato e non è in alcun modo collegata con rappresentazioni interne o intenzioni degli agenti. T.C. Schelling, *Dynamic models of segregation*, in *Journal of mathematical sociology*, 1971, 1 (2), pp. 143-186.

- 16 Introdotti negli anni Cinquanta del secolo scorso da John von Neumann e Stanislaw Ulam, gli automi cellulari sono modelli utilizzati nello studio dell'evoluzione di sistemi complessi. L'idea di fondo di questi modelli è quella di descrivere il comportamento di un sistema complesso "dal basso", simulandone l'evoluzione mediante interazioni fra elementi, detti "celle", che seguono semplici regole. Un automa cellulare è perciò composto da un reticolo (griglia) di celle; ciascuna cella assume stati (colori) differenti in funzione di leggi matematiche che ne regolano l'evoluzione in relazione al comportamento delle celle vicine. Per un approfondimento, si veda S. Albeverio, P. Giordano, A. Vancheri, *Teoria generale degli automi cellulari*, in *Metodi e modelli matematici per le dinamiche urbane*, a cura di S. Albeverio, P. Giordano, A. Vancheri, Milano, Springer, 2021, pp. 33-46.
- 17 Si v., a titolo esemplificativo, J. Uyheng, D. Bellutta, K.M. Carley, *Bots amplify and redirect hate speech in online discourse about racism during the COVID-19 pandemic*, in *Social media and society*, 2022, 8 (3); S.U. Noble, *Algorithms of oppression*, New York, New York University Press, 2018; X. Qiu, D.F. Oliveira, A.S. Shirazi, A. Flammini, F. Menczer, *Limited individual attention and online virality of low-quality information*, in *Nature human behaviour*, 2017, 1 (7), pp. 1-7; M. Del Vicario, G. Vivaldo, A. Bessi, F. Zollo, A. Scala, G. Caldarelli, W. Quattrociocchi, *Echo chambers: Emotional contagion and group polarization on facebook*, in *Scientific reports*, 2016, 6 (1), pp. 1-12.
- 18 Seppur relativo a un diverso tema di indagine, merita di essere segnalato un interessante studio sul legame che unisce la topologia delle reti interazione l'emergere di norme sociali di correttezza e cooperazione. Sul punto, si veda M. Mosleh, B. Heydari, *Fair topologies: Community structures and network hubs drive emergence of fairness norms*, in *Scientific reports*, 2017, 7 (1), pp. 1-9.

e implicazioni giuridiche delle nuove forme di discriminazione e, allo stesso tempo, per predisporre contromisure adeguate alla particolare natura delle interazioni che hanno luogo nella società digitalizzata. Dopotutto, è evidente come non si possano porre rimedi giuridici a dinamiche ignote, non si possono contrastare violazioni di diritti tutelati dall'ordinamento in assenza di strumenti di analisi della dimensione tecnica degli algoritmi e dei processi da essi innescati nel loro concreto operare.

La questione emerge in modo chiaro in un recente studio del *Panel for the Future of Science and Technology* del Parlamento europeo che analizza le sfide da affrontare sul piano tecnico e normativo per orientare gli algoritmi¹⁹ a fini socialmente desiderabili. In quest'ottica, l'utilizzo di sistemi di *algorithmic decision making* (ADS) viene subordinato al rispetto di criteri che includono non solo l'equità, l'imparzialità e la non discriminatorietà²⁰ ma anche - per quanto di nostro specifico interesse - la "comprensibilità" del codice informatico ossia la possibilità concreta di ricostruire i nessi che collegano l'input (i dati) e l'output (le scelte) dei sistemi decisionali²¹. Non è un caso che in più occasioni²² la Commissione europea abbia identificato nell'audit e nelle valutazioni di conformità ai criteri sopra elencati, gli strumenti chiave per garantire l'affidabilità degli algoritmi e promuovere la fiducia nei loro confronti²³.

I fattori che si oppongono alla implementazione di questi strumenti sono tuttavia numerosi. Se per un verso mancano ancora, sul piano normativo e istituzionale, le condizioni necessarie a creare meccanismi pubblici di analisi sistematica degli algoritmi²⁴, per altro verso resta da affrontare il tema

19 Sui sistemi di decisione algoritmica si v. C. Castelluccia, D. Le Métayer, *Understanding algorithmic decision-making: Opportunities and challenges*. *European Parliamentary Research, Service*, 2019.

20 Il rapporto parla a tale proposito di "intrinsic requirements" e definisce l'equità come "assenza di parzialità indesiderabile" e la "discriminazione" come una forma particolare di ingiustizia legata all'uso di tipi specifici di dati (come origine etnica, opinioni politiche, genere, ecc.).

21 Lo studio articola la comprensibilità (*understandability*) nei due concetti di *transparency* ed *explainability*. La prima viene identificata con la messa a disposizione del pubblico (o di autorità competenti di audit e certificazione) di una serie di informazioni: il codice informatico alla base dell'ADS; la documentazione di progettazione; i dataset e i parametri di apprendimento nei casi in cui l'ADS si basa su sistemi di *machine learning*. La *explainability* è definita come la disponibilità di spiegazioni sull'ADS. Contrariamente alla trasparenza, la spiegabilità richiede la consegna di informazioni oltre l'ADS stesso. Le spiegazioni possono essere di diversi tipi (operative, logiche o causali); globali (dell'intero algoritmo) o locali (su risultati specifici); e possono assumere forme diverse (alberi decisionali, istogrammi, evidenziazioni di immagini o testo, esempi, controesempi, ecc.). I punti di forza e di debolezza di ciascuna modalità di spiegazione dovrebbero essere valutati in relazione ai destinatari (ad esempio professionale o individuale), al loro livello di competenza e ai loro obiettivi (per contestare una decisione, intraprendere azioni per ottenere una decisione, verificare la conformità con obblighi legali, ecc.).

22 In questo senso, si veda lo European Commission white paper on Artificial Intelligence – A European approach to excellence and trust (COM(2020) 65 final), 2020. Nel documento si legge: "testing centres should enable the independent audit and assessment of AI-systems [...]. Independent assessment will increase trust and ensures objectivity. It could also facilitate the work of relevant competent authorities. Il tema della trasparenza del decision-making algoritmico è al centro di un numero crescente di lavori interdisciplinari. Tra questi, si veda B. Lepri, N. Oliver, E. Letouzé, A. Pentland, P. Vinck, *Fair, transparent, and accountable algorithmic decision-making processes*, in *Philosophy & Technology*, 2018, 31 (4), pp. 611-627.

23 Il *white paper* ipotizza a tal fine la costituzione di organismi designati dagli Stati membri in grado di garantire, al contempo, competenze tecniche, indipendenza e rappresentanza delle diverse categorie di soggetti interessati dei processi di evoluzione algoritmica in atto.

24 Sul punto si veda J. Danaher, M.J. Hogan, C. Noone et al. *Algorithmic governance: Developing a research agenda through the power of collective intelligence*, in *Big data & Society*, 2017, 4 (2). Gli autori presentano i risultati di un workshop che ha riunito un gruppo multidisciplinare di studiosi per considerare: a) gli ostacoli ad un uso legittimo ed efficace della *governance* algoritmica e b) i metodi di ricerca necessari per affrontare gli ostacoli messi a fuoco. Dal lavoro emerge come gli ostacoli che impediscono risposte adeguate ai rischi insiti negli algoritmi siano molteplici. Accanto all'opacità degli algoritmi sopra

nodale dell'opacità delle macchine algoritmiche e delle diverse forme attraverso cui si materializza la loro idoneità lesiva. La natura ibrida - al tempo stesso tecnica, giuridica, sociologica ed economica - delle analisi da compiere richiede la messa a punto di programmi di ricerca adeguati alle caratteristiche del fenomeno oggetto di indagine programmi, a ben vedere, ancora in via di definizione²⁵.

Un paradigma meritevole di attenzione in questa prospettiva è senza dubbio rappresentato dai c.d. *Critical data and algorithm studies*²⁶ (CDS), area di studi emersa nel corso degli ultimi anni anche in ragione dell'impatto avuto, sull'opinione pubblica e sulla coscienza collettiva, da casi di cronaca come quello di *Cambridge Analytica*. Il tema di ricerca non è nuovo in senso assoluto. La necessità di una riflessione critica sui Big data e sui sistemi di *algorithmic decision making* è stata già in passato rimarcata da *computer scientist* particolarmente attenti ai risvolti negativi potenzialmente connessi alle tecnologie ed alle euristiche frutto delle loro ricerche²⁷. La riflessione maturata sul punto, tuttavia, era stata sino ad oggi caratterizzata da una focalizzazione quasi esclusiva sugli aspetti tecnici e da uno sguardo orientato alle soluzioni pratiche senza la necessaria integrazione tra la prospettiva tecnica e quella *lato sensu* critica.

Nel tentativo di rispondere a questa carenza, i CDS propongono un nuovo approccio allo studio critico delle pratiche digitali e alle loro implicazioni, puntando all'integrazione di discipline diverse dal diritto, all'informatica, dalla sociologia, all'economia, dalle scienze comportamentali all'etica. Le questioni da affrontare, d'altra parte, richiedono chiavi di lettura e conoscenze che trascendono i confini di singole aree di ricerca: In che modo i sistemi algoritmici finiscono con l'incorporare valori e pregiudizi e come è possibile rendersene conto? Come possiamo operationalizzare l'equità, pensare all'etica nei sistemi algoritmici e gestire la responsabilità in processi decisionali plasmati - in modo consapevole o anche inconsapevole - da sviluppatori, utenti, manager, leggi, codice informatico, tecnologie hardware? Attraverso quali percorsi la cultura, le regole giuridiche e la società influenzano la creazione di algoritmi e viceversa?

In linea con l'ampiezza di queste domande, l'orizzonte metodologico dei CDS è estremamente eterogeneo e include -con modalità ancora in larga parte da definire- pratiche di ricerca molto diverse. Una recente nota definizione dell'agenda di ricerca dei CDS²⁸ ne elenca ben diciannove, passando dalle *survey* etnografiche, all'*action-research*, dalle analisi forensi alle perizie informatiche, dallo studio dei casi giudiziari alla raccolta di dati secondari. La scelta non deve sembrare eccessiva. L'indagine critica degli algoritmi e delle loro possibili aberrazioni presuppone la ricostruzione di una trama di relazioni - quella che lega dati, algoritmi, dinamiche economiche e sociali, rapporti giuridici - intricata e

evocata, gli autori elencano numerosi altri problemi ugualmente bisognosi di risposta che vanno dal tecno-ottimismo acritico dei politici e degli stakeholder tecnologici, alla mancanza di cultura e competenza tecnica tra politici e operatori del settore pubblico.

- 25 Per una riflessione sul tema, sia consentito rinviare a N. Lettieri, *Antigone e gli algoritmi. Appunti per un approccio giusfilosofico*. Modena, Mucchi, 2020.
- 26 Alcuni autori distinguono *Critical data* e *Critical algorithm studies* utilizzando come criterio discretivo la focalizzazione degli studiosi sui dati o sugli algoritmi. Una serie di fattori (la sostanziale identità di obiettivi, lo stretto legame esistente tra dati e algoritmi, le caratteristiche metodologiche) ci inducono a trattare le due linee di pensiero in maniera unitaria. Caratterizzata da una natura fortemente interdisciplinare, la letteratura riconducibile all'area dei nascenti *critical data and algorithm studies* è già ricca di contributi interessanti spesso scritti, a più mani, da studiosi di diversa estrazione. Per un inquadramento dell'arte si veda R. Kitchin, *Thinking critically about and researching algorithms*, in *Information, communication & society*, 2017, 20 (1), pp. 14-29.; A. Iliads, F. Russo, *Critical data studies: An introduction*, in *Big Data & Society*, 2016, 3 (2). Un esempio paradigmatico di ricerche riconducibili al filone dei CDS è offerto da S. Noble, *Algorithms of oppression*, cit.
- 27 Si veda J. Isaak, M.J. Hanna, *User data privacy: Facebook, Cambridge Analytica, and privacy protection*, in *Computer*, 2018, 51 (8), pp. 56-59.
- 28 J. Danaher, M.J. Hogan, C. Noone, et al, *Algorithmic governance*, op. cit.

tecnicamente complessa. L'elettismo metodologico e la cross-fertilizzazione tra aree di ricerca rappresentano una scelta in qualche modo obbligata, al di là del contesto fattuale preso in considerazione.²⁹

Un paradigma cui guardare con attenzione in questa prospettiva è senza dubbio quello delle scienze sociali computazionali³⁰, una delle aree di ricerca che, insieme all'economia comportamentale e alla psicologia economica³¹, sta aprendo le prospettive più interessanti alla comprensione dei fenomeni sociali. Frutto dell'incontro³² tra le domande di ricerca delle scienze sociali, il quadro di teorico offerto dalla teoria della complessità, i metodi della scienza computazionale e il diluvio di dati che caratterizza la società dell'informazione³³, la scienza sociale computazionale può essere definita, per usare le parole di uno dei suoi principali rappresentanti, come "*the interdisciplinary investigation of the social universe on many scales, ranging from individual actors to the largest groupings, through the medium of computation*"³⁴. Combinando teorie sociali, computazione, dati ed esperimenti in una inedita visione sistemica e multidisciplinare, la scienza sociale computazionale sembra poter illuminare i processi alla base della complessità sociale inclusa l'interazione tra scelte individuali, dinamiche collettive e fattori tecnologici che, come visto, condiziona la discriminazione algoritmica.

L'apparato metodologico delle scienze sociali computazionali, d'altra parte, include oggi una varietà di approcci e pratiche di ricerca³⁵ che, specie quando combinati tra di loro, hanno mostrato di essere estremamente potenti in tal senso. La *network analysis*³⁶, per fare un esempio, offre nuove opportunità per analizzare in termini quantitativi come le caratteristiche di strutture di interazione in senso lato intese (reti sociali, biologiche, finanziarie, tecnologiche) incidano sul comportamento di un sistema (una rete criminale, un sistema di alleanze, una rete di computer) e delle sue componenti.

-
- 29 P. Ball, *Why society is a complex matter: Meeting twenty-first century challenges with a new kind of science*, 2012 Springer Science & Business Media; P. Ormerod, *Networks and the need for a new approach to policymaking*, in: T. Dolphin, D. Nash (a cura di), *Complex New World. Translating New Economic Thinking into Public Policy*, Institute for Policy Research, 2012, London, pp. 28-38; J.D. Farmer, D. Foley, *The economy needs agent-based modelling*, in *Nature*, 2009, 460(7256), pp. 685-686.
- 30 Per un'introduzione al tema, si v. C. Cioffi-Revilla, *Introduction to computational social science*, London and Heidelberg: Springer, 2014; R. Conte, N. Gilbert, N., G. Bonelli, et al., *Manifesto of computational social science. The European Physical Journal Special Topics*, 2012, 214(1), 325-346; G.N. Gilbert (Ed.), *Computational social science*, 2010 (Vol. 21), Sage.
- 31 D. Kahneman, *Maps of bounded rationality: Psychology for behavioral economics*, in *American economic review*, 2003, 93(5), pp. 1449-1475.
- 32 P. Ormerod, *N squared: Public policy and the power of networks*. RSA Pamphlet, 2010.
- 33 C. Anderson, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, 2008, 16 Wired Magazine 7.
- 34 C. Cioffi-Revilla, *Introduction to Computational Social Science*, cit., p. 2.
- 35 Il paradigma delle scienze sociali computazionali comprende una varietà in evoluzione di metodologie di ricerca: estrazione automatica di informazioni, modelli di complessità, simulazione sociale, analisi dei social network e analisi geospaziale. Per un approfondimento, si v. C. Cioffi-Revilla, *Introduction to computational social science*, cit.
- 36 La *network analysis* è connessa alla *network science*, ambito di ricerca che studia reti complesse in senso lato intese (dalle reti informatiche alle reti biologiche) rappresentando l'oggetto della propria indagine come una struttura (grafo o *network*) costituita da nodi (o vertici) e connessioni tra questi (archi). Per far questo, la *network science* attinge a una pluralità di teorie e metodi provenienti da aree disciplinari diverse: la *teoria dei grafi* (dalla matematica), la *meccanica statistica* (dalla fisica), l'estrazione di dati e la visualizzazione delle informazioni (dall'informatica), la modellazione inferenziale (dalla statistica). Nelle scienze sociali, la *network science* ha alimentato un filone di ricerca sull'influenza della relazione sociale sul comportamento collettivo offrendo opportunità per illuminare proprietà dei sistemi studiati (resilienza, vulnerabilità, scomponibilità, funzionalità) difficili da ricavare attraverso metodi tradizionali. Le tecniche delle *network analysis* hanno iniziato a trovare applicazione anche in ambito giuridico per scopi diversi che vanno dall'analisi delle proprietà dei corpora giuridici allo studio delle reti criminali. Per un'introduzione alla *social network analysis*, si veda S. Wasserman, K. Faust, *Social network analysis: Methods and applications*. Cambridge, Cambridge University Press, 1994. Per una panoramica sui possibili impieghi in ambito giuridico, sia consentito rinviare a R. Winkels, N. Lettieri, S. Faro, *Network analysis in law*, ESI, Napoli, 2013.

Un discorso analogo può essere fatto per le tecniche di simulazione sociale, in particolare la simulazione basata su agente (*Agent-Based Modeling - ABM*)³⁷, che consente oggi di studiare sperimentalmente l'emersione e l'evoluzione di fenomeni macroscopici e collettivi (es. il diffondersi di norme sociali o di forme di segregazione razziale) a partire dalla descrizione del comportamento individuale di singole entità distinte tra loro.

Alla luce di quanto detto, appare evidente che le dinamiche alla base della discriminazione algoritmica hanno di fronte a sé nuove opportunità per essere comprese e analizzate criticamente. Il percorso che conduce in questa direzione, resta tuttavia ancora largamente da esplorare.

4. **Computational law: gli algoritmi da problema a soluzione**

La migliore comprensione dei processi che favoriscono direttamente o indirettamente fenomeni di discriminazione mediata dalle macchine non rappresenta l'unica direzione verso cui guardare per contrastare gli effetti perversi dell'evoluzione tecnologica. All'attività di studio e analisi critica occorre affiancare, per altro verso, la ricerca di contromisure adeguate a contesti in cui l'interazione sociale ed economica è interamente mediata dalle tecnologie dell'informazione e l'efficacia dei rimedi giuridici tradizionali è obiettivamente limitata.

Nelle piattaforme di lavoro o sui *social media*, a ben vedere, il confine tra ciò che può o non può essere fatto, è determinato non solo dalla norma giuridica astrattamente applicabile ma anche, in maniera più diretta, dall'infrastruttura tecnologica e dal codice informatico che abilitano e governano le nostre interazioni. In scenari di questo tipo, il rischio che le norme di legge risultino inadeguate sul piano della tempestività e della stessa capacità di imporsi è molto concreto.

Un orizzonte meritevole di esplorazione, in quest'ottica, è rappresentato dalla prospettiva di un uso regolativo degli algoritmi in senso lato intesi, di un impiego di strumenti hardware e software per dettare norme, implementare la tutela di interessi giuridicamente rilevanti e contrastare la violazione di norme di diritto.

L'idea, variamente declinata oggi nei paradigmi della "tecnoregolazione" e della "*computational law*", non è nuova³⁸. Le sue radici risalgono agli anni Novanta, quando Internet ha fatto gradualmente

37 Basata sull'identificazione della spiegazione scientifica con la riproduzione "*in silico*" (cioè in una simulazione al computer) di processi sociali oggetto di indagine, la simulazione basata su agenti si pone alla base di un approccio generativo alla ricerca al cui interno le dinamiche e le strutture sociali di livello macro sono interpretate, descritte, riprodotte e, quindi, spiegate, come il risultato delle interazioni di livello micro tra entità computazionali (agenti) che simulano il comportamento di individui reali. Alla luce di queste premesse, una simulazione ad agenti include tipicamente un insieme di attori/agenti, un insieme di regole di interazione e un ambiente di cui sono state definite le caratteristiche dinamiche, organizzative e spaziali. Nel corso degli anni, la simulazione sociale basata su agenti ha mostrato di essere particolarmente adatta a investigare di dinamiche socioeconomiche che sono difficili da trattare con approcci di tipo statistico. Grazie a questa caratteristica, le simulazioni non solo si sono diffuse nell'ambito della scienza ma hanno anche richiamato l'attenzione delle istituzioni coinvolte nel *policy making*. Per un'introduzione alla scienza sociale generativa e alla simulazione ad agente, si veda, in una letteratura ormai ampia, N. Gilbert, *Agent-based models*, London, Sage, 2008; J.M. Epstein, *Generative social science*, Princeton and Oxford, Princeton University Press, 2012. Sulle intersezioni tra simulazione sociale e diritto, sia consentito rinviare a N. Lettieri, *Ius in silico. Diritto computazione simulazione*, Napoli, ESI, 2013; N. Lettieri, D. Parisi, *Neminem laedere. An evolutionary agent-based model of the interplay between punishment and damaging behaviours*, in *Artificial intelligence and law*, 2013, 21(4), pp. 425-453.

38 Il tema è oggetto di una riflessione ormai molto ampia e di livello internazionale. Tra i primi contributi in materia si veda R. Brownsword, *What the world needs now: Techno-regulation, human rights and human dignity*, in *Human rights*, Hart Projected

emergere la capacità del codice informatico di affiancare la legge dello stato nello stabilire quello che le persone possono o non possono fare. Suscettibili di applicazione automatica, le regole inscritte nei *software*, negli standard tecnici e nelle scelte di design che governano il cyberspazio hanno infatti presto mostrato come gli algoritmi diano corpo a un nuovo tipo di sistema normativo in grado di aprire nuove frontiere alla regolazione della vita sociale.

La "*Lex Informatica*"³⁹ che domina web è stata, tuttavia, solo il punto di partenza. La presa di coscienza della funzione intrinsecamente regolativa delle ICT si è rapidamente tradotta nell'idea di servirsi in maniera intenzionale di artefatti tecnologici per implementare la disciplina giuridica di rapporti economici e sociali. È stato guardando in questa direzione che Roger Brownsword ha parlato per la prima volta, negli anni 2000, di "*normative environment*"⁴⁰ prefigurando infrastrutture tecnologiche capaci di implementare in modi nuovi la funzione ordinamentale solitamente assolta dalle norme di diritto.

In contesti come il web, la tecnologia è in effetti capace di imporre vincoli comportamentali caratterizzati da un livello di incisività e immediatezza molto superiori rispetto a quello delle disposizioni di legge. Mentre queste ultime si limitano a definire precetti, la prima può determinare in maniera diretta le azioni che gli individui possono o non possono compiere, rendendo superfluo l'intervento di una autorità di controllo che punisca, *ex post*, i trasgressori⁴¹.

Lo sviluppo scientifico e tecnologico svolge, in questa prospettiva, un ruolo determinante consentendo la sperimentazione di strategie regolative di tipo diverso. Alle tecnologie utilizzate nei primi sistemi di *Digital Rights Management*⁴², forma embrionale di tecnodiritto, se ne sono affiancate altre in

Date, 2004, p. 203 ss. Per un'analisi organica e più recente si veda M. Hildebrandt, J. Gaakeer (eds.), *Human law and computer law: Comparative perspectives*, Heidelberg, Springer, 2013. Il volume pone la sua attenzione non solo sull'applicazione automatizzata delle norme legali, ma estende il concetto di tecno-regolazione riconducendo ad esso la progettazione di soluzioni tecniche volte a influenzare i comportamenti promuovendo la conformità alle norme attraverso avvertenze o suggerimenti mirati. In tal senso vedi B. Van Den Berg, R.E. Leenes, *Abort, retry, fail: Scoping techno-regulation and other techno-effects*, ivi, pp. 67-87. Sullo stesso punto M. Hildebrandt, *Legal protection by design: Objections and refutations*, in *Legisprudence*, 2011, 5, pp. 223-248; R. Leenes R., *Framing techno-regulation: An exploration of state and non-state regulation by technology*, in *Legisprudence*, 2011, 5, pp. 143-169. In italiano si veda, *ex multis*, A.C. Amato Mangiameli, *Tecno-diritto e tecno-regolazione. Spunti di riflessione*, in *Rivista di filosofia del diritto*, 2017, 6 (Speciale), pp. 87-112.

39 Si veda J.R. Reidenberg, *Lex informatica: The formulation of information policy rules through technology*, in *Texas Law Review*, 1997, 76, p. 553 ss. L'intuizione viene ripresa e resa in termini ancor più radicali da Lawrence Lessig dello Stanford Center for Internet Society, con la nota affermazione "in the cyberspace, the code is law", cfr. L. Lessig, L. *Code and other laws of cyberspace*, New York, Basic Books, 2009.

40 Un esempio di "normative environment" viene collocato da Brownsword al di fuori del web. L'autore cita un'architettura tecnologica sviluppata dai titolari un golf club del Warwickshire per impedire furti ed usi impropri dei propri kart. Per far fronte a questa evenienza, il club ha installato su ogni kart un sistema di controllo automatico basato su GPS in grado di per dar vita a una sorta di "una geo-recinzione virtuale intorno al club". Grazie a questa forma tecnologica di tutela del diritto di proprietà, i titolari del club possono non impedire i furti, ma anche un uso dei kart in aree bagnate o oggetto di manutenzione. Chiunque dovesse provare a violare la geo-recinzione scoprirebbe che il kart si blocca automaticamente. R. Brownsword, *What the world needs now*, cit.

41 In tal senso, si veda S. Hassan, P. De Filippi, *The expansion of algorithmic governance: from code is law to law is code. Field actions science reports*, in *The journal of field actions*, 2017, Special Issue 17, pp. 88-90.

42 I *digital right management systems* (DRMs) sono sistemi tecnologici sviluppati a partire dai primi anni Duemila mediante i quali i titolari di diritto d'autore (e dei diritti connessi) possono tutelare, esercitare ed amministrare i propri diritti nell'ambiente digitale (ad esempio, limitando il numero di copie possibili di un brano in formato digitale). Sul punto si veda Q. Liu, R. Safavi-Naini, N.P. Sheppard, *Digital rights management for content distribution*, in *Proceedings of the Australasian information security workshop conference on ACSW frontiers*, 2003, 21, pp. 49-58.

grado di abilitare l'emersione di modelli tecnoregolativi applicabili a contesti nuovi e sempre più complessi. La *blockchain*⁴³, architettura che sfrutta database distribuiti e criptati per gestire in modo sicuro le transazioni che avvengono tra i nodi di una rete, rappresenta, da questo punto di vista, un esempio paradigmatico. Dopo le prime applicazioni legate alla creazione di valute digitali, essa ha rapidamente trovato spazio in ambito giuridico grazie alla possibilità di caricare al suo interno brani di codice - gli "*smart contract*" alla base del modello della "*computational law*"⁴⁴ che consentono di affidare a un software la costituzione ed esecuzione degli obblighi derivanti da un rapporto negoziale a prescindere dalla stipula di un contratto tradizionale.

L'elenco potrebbe continuare. La riflessione sui modelli di tecnoregolazione ha condotto a un numero considerevole di esperienze e ipotesi sperimentali che vanno dall'uso della crittografia e di piattaforme web per tutelare la proprietà intellettuale nella circolazione dei modelli industriali 3D, allo sviluppo di agenti intelligenti in grado di assistere gli utenti durante le loro interazioni *on line* per proteggerli dalla minaccia di truffe⁴⁵.

Come è stato da più parti evidenziato⁴⁶, ci troviamo di fronte all'emergere di una nuova dimensione della normatività, alla trasposizione sul piano tecnologico di tutte le attività che integrano il processo regolativo "ordinario" dalla identificazione di soggetti e condotte tipici di una data fattispecie, all'esecuzione dei precetti passando attraverso la promozione di comportamenti conformi alla legge mediante avvertenze o suggerimenti mirati riconducibili alla prospettiva del *nudge*⁴⁷. Considerata la sua forte connotazione tecnologica, la discriminazione algoritmica non può rimanere estranea a tutto ciò.

-
- 43 Grazie alle caratteristiche di trasparenza e incorruttibilità delle registrazioni operate al suo interno, la *blockchain* consente di generare valore e fiducia senza la necessità di rapporti fiduciari diretti tra le parti che interagiscono, consentendo così il coordinamento di un gran numero di individui che non si conoscono (e quindi non si fidano necessariamente l'uno dell'altro). Per un'introduzione alla *blockchain* letta nelle sue proiezioni giuridiche si veda, P. De Filippi, *Blockchain and the law*, in P. De Filippi, A. Wright, *Blockchain and the Law. The rule of code*. Harvard University Press, 2018.
- 44 L'espressione "*computational law*" allude oggi a un diritto fatto di leggi e contratti scritti direttamente nel codice del computer e capaci di esecuzione automatica al verificarsi delle condizioni prestabilite. L'integrazione del diritto computazionale e delle tecno regole nella pratica giuridica ordinaria è una realtà destinata a trasformare la natura stessa del diritto come sostenuto da Mireille Hildebrandt, uno dei giuristi più attenti all'evoluzione in senso computazionale della normatività giuridica responsabile del progetto CoHuBiCoL - *Counting as a human Being in the era of computational law* (<https://www.cohubicol.com>), che investe proprio questi temi.
- 45 Per una disamina delle esperienze citate, si veda T. Kerikmäe, A. Rull (eds.), *The future of law and eTechnologies*, Vol. 3, Springer International Publishing, 2016.
- 46 Sul punto si veda tra gli altri, C. Sarra, "*Iper-positività*": la riduzione del giuridicamente lecito al tecnicamente possibile nella società dell'informazione, in *Positività giuridica: Studi ed attualizzazioni di un concetto complesso*, a cura di C. Sarra e M.a Isabel Garrido Gómez, Padova University Press, 2018, pp. 95-125. Come sottolineato dall'autore, l'informatizzazione della società sta realizzando una transizione che coinvolge "il concetto stesso di diritto quale lo si è inteso dalla modernità ai giorni nostri, e ciò esattamente perché ambisce a realizzare una forma del giuridico che - sempre più conformata all'automatismo informatico - non ha bisogno della volontà per realizzarsi nel mondo". Per dirla con le parole di Giovanni Sartor, ci si muove verso un modello di diritto nel quale a ciascun individuo è «lecito fare quanto egli di fatto è in grado di fare [...] ma egli è in grado di fare solo ciò che è stato abilitato a fare», cfr. G. Sartor *Internet e il diritto*, in *Temi di diritto dell'informatica*, G. Sartor, C. Di Cocco (a cura di), Torino, 2017, pp. 4-5.
- 47 La riflessione teorica in tema di tecnoregolazione si è soffermata di recente sugli usi regolativi della tecnologia orientati non ad imporre quanto ad influenzare il comportamento umano attraverso forme di persuasione o "nudge", la spinta gentile teorizzata in R.H. Thaler, C.R. Sunstein, *Nudge: Improving decisions about health, wealth, and happiness*, New York, Penguin, 2009. Sul punto si veda B.V.D. Berg, R.E. Leenes, *Abort, retry, fail: scoping techno-regulation and other techno-effects*, in *Human law and computer law: comparative perspectives*, edited by M. Hildebrandt, J. Gaaker, Dordrecht, 2013, pp. 67-87.

5. Algoritmi contro la discriminazione algoritmica: una sperimentazione

Il quadro delineato nelle pagine precedenti ha lasciato chiaramente intravedere come la messa a punto di strategie di risposta ai rischi della discriminazione algoritmica richieda per ragioni diverse di rivolgere lo sguardo verso le scienze e alle tecnologie dell'informazione. Il confronto, anche di tipo operativo, con la *computer science* e i suoi strumenti, con la raccolta e l'analisi dei dati, si impone non solo per comprendere meglio i processi di tipo discriminatorio influenzati dagli algoritmi, ma anche per predisporre nuove e più efficaci forme di tutela.

In quest'ottica si muove un'attività di ricerca interdisciplinare che punta a esplorare sperimentalmente⁴⁸ nuovi approcci di analisi critica e regolazione in un contesto, quello della *gig economy*, caratterizzato da una serie di questioni - atomizzazione dei processi lavorativi, controllo pervasivo e occulto dei lavoratori, scarsità di dati e informazioni indipendenti- tra cui figurano anche le nuove forme di discriminazione e ingiustizia mediata dagli algoritmi.

Il punto di partenza è una proposta - già formalizzata sul piano teorico⁴⁹ - che *specula* sulla possibilità di raggiungere livelli più elevati di protezione per i lavoratori della *gig-economy* integrando rimedi legali e strumenti tecnologici. Sulla base di questa premessa, l'indagine punta a definire una prospettiva di ricerca e un quadro metodologico funzionali non solo ad una migliore comprensione

48 Lungo questa linea di ricerca si muovono anche altre sperimentazioni condotte in collaborazione con la prof.ssa Delfina Malandrino, il prof. Rocco Zaccagnino e il dott. Alfonso Guarino del Dipartimento di Informatica con cui è da tempo attiva una collaborazione intorno alle intersezioni fra diritto e *computer science*. Sul tema della tecnoregolazione, si veda, fra i più recenti, N. Lettieri, D. Malandrino, R. Zaccagnino, A. Guarino, *The eye of the rider. Visualization and data-driven heuristics for the critical analysis of gig economy*, in *26th International Conference Information Visualisation*, (IV) 2022 Ebad Banissi, pp. 379-384; A. Guarino, N. Lettieri, D. Malandrino, R. Zaccagnino, *A machine learning-based approach to identify unlawful practices in online terms of service: analysis, implementation and evaluation*, in *Neural computing and applications*, 2021, 33 (24), pp. 17569-17587; R. Zaccagnino, C. Capo, A. Guarino, N. Lettieri, D. Malandrino, *Techno-regulation and intelligent safeguards*, in *Multi-media tools and applications*, 2021, 80 (10), pp. 15803-15824.

49 Si veda, con specifico riferimento alla tutela dei lavoratori della *gig economy*, M. De Minicis, S. Donà, N. Lettieri, M. Marocco, *Disciplina e tutela del lavoro nelle digital labour platform. Un modello di tecnoregolazione*, in *Economia & lavoro*, 2019, 53 (3), pp. 89-106. La proposta ruota intorno al "Sistema Integrato di Crowdwork" (SIC), infrastruttura pubblico-privata pensata per favorire il controllo dello stato sul mercato del *crowdwork* attraverso un *layer* tecnologico destinato a mediare le interazioni che hanno luogo tra tutti i soggetti a vario titolo coinvolti nelle dinamiche di mercato del lavoro digitale: amministrazioni pubbliche (es. ministeri, ispettorati del lavoro, istituti di previdenza sociale), *crowdworker*; DLP; altri attori (es. associazioni sindacali e sindacali, potenziali clienti e crowd workers, cittadini comuni). In termini generali, l'infrastruttura dovrebbe supportare più funzioni collegate fra loro: i) registrazione dei rapporti lavorativi tra *labour platform* e *crowdworker* a fini di garanzia della certezza e trasparenza delle operazioni; ii) raccolta e trasmissione alle Pubbliche Amministrazioni dei dati relativi alle interazioni tra lavoratori e piattaforme; iii) stipula di smart contract; iv) conservazione e gestione dei dati di tipo economico e previdenziale relativi alle prestazioni occasionali svolte dal lavoratore; v) implementazione di meccanismi reputazionali orientati a condizionare l'operato delle piattaforme e a proteggere i *crowdworker*. Per implementare queste funzioni, l'architettura del SIC contempla due moduli. Un primo modulo (*Reputation and transparency module*), parzialmente realizzato in via prototipale, punta a implementare meccanismi di reputazione e l'accesso del pubblico alle informazioni sul mercato del lavoro collettivo attraverso un'app mobile (pensata per consentire ai lavoratori di esprimere le proprie valutazioni sull'operato delle piattaforme digitali) e un sito web (per consentire l'accesso del pubblico ai dati generati dall'analisi dei dati prodotti dai *crowdworker* attraverso l'interazione con il SIC). Un secondo modulo in via di sviluppo (*Social Security and Contracts Module*) dovrebbe sfruttare la tecnologia *blockchain* e *smart contract* per abilitare e tenere traccia di qualsiasi interazione che ha luogo nel mercato del lavoro digitale.

delle dinamiche del lavoro digitale, ma anche ad una *governance* più efficace dei rapporti economici e giuridici che avvengono al suo interno.

La sperimentazione ruota intorno a *GigAdvisor*, prototipo di una piattaforma accessibile via *web* e *app mobile* concepita come “laboratorio virtuale” per indagini collegate tanto al paradigma dei *Critical data studies* e della *computational social science*, quanto a quello della tecnoregolazione⁵⁰. La componente principale del sistema è un modulo che permette ai lavoratori della *gig-economy* - previa registrazione - di condividere e discutere valutazioni relative alle esperienze avute con le piattaforme (es. *Amazon Mechanical Turk*, *Deliveroo*, *Uber*) per cui lavorano.

Le informazioni raccolte attraverso questa semplice forma di interazione sono il punto di partenza di analisi e esperimenti destinati a coinvolgere diverse categorie di soggetti. Da una parte ci sono gli utenti “esterni” (lavoratori delle piattaforme, cittadini o anche, in prospettiva, pubbliche amministrazioni) che si servono della piattaforma per scopi di carattere pratico (esprimere valutazioni, discutere, osservare il *rating* delle piattaforme per decidere, ad esempio, con chi lavorare o, ancora per svolgere attività di controllo indipendente). Dall'altra, ci sono i ricercatori interessati a esplorare le dinamiche del lavoro digitale (inclusi fenomeni di tipo discriminatorio), utilizzando le funzionalità di *data analytics* messe a disposizione dalla piattaforma ovvero scaricando i dati per analizzarli con altri strumenti.

Una volta integrate con dati di tipo spaziale e temporale e analizzate con i metodi delle scienze sociali computazionali dall'*information extraction* al social GIS, le valutazioni espresse dai *crowdworker* possono tradursi in conoscenze in vario modo funzionali alla lotta alla discriminazione algoritmica. Si pensi in questo senso alla possibilità di indagare su nuove basi empiriche l'effettivo rispetto delle norme antidiscriminatorie attraverso l'analisi dettagliata delle condizioni di lavoro o attraverso l'identificazione, nel comportamento delle *digital labour platform*, di *pattern* ricorrenti indici di discriminazioni e distorsioni indotte dai meccanismi di *governance* algoritmica della forza lavoro.

Accanto alle attività di studio appena delineate, il progetto punta anche a favorire l'esplorazione di soluzioni tecnoregolative in vario modo funzionali anche a contrastare la discriminazione. Una prima strada, già in parte esplorata sperimentalmente nello sviluppo di *GigAdvisor*, punta a sfruttare il meccanismo reputazionale (basato su valutazioni e dati relativi anche a dinamiche discriminatorie) per spingere le *labor platform* verso l'adozione di meccanismi di *workforce analytics* e *workforce management* più equi sfruttando il potere condizionante dell'immagine reputazionale ricavabile, anche attraverso tecniche di *data-analytics*, dalle valutazioni espresse dai lavoratori⁵¹. Una seconda strada, in via di sperimentazione, ruota intorno all'ipotesi di un utilizzo della *blockchain* e degli *smart contract* per garantire la conformità dei contratti lavoratore-piattaforma alle norme poste dall'ordinamento vigente a tutela dei diritti dei lavoratori incluso quello alla non discriminazione.

50 Sul punto: N. Lettieri, A. Guarino, D. Malandrino, R. Zaccagnino, *Platform economy and techno-regulation – experimenting with reputation and nudge*, in *Future Internet*, 2019, 11 (7), p. 163, <https://bit.ly/2WXEKQp>.

51 Nella storia recente della *gig economy* non mancano evidenze in merito a come l'assegnazione di rating possa spingere le piattaforme a modificare il loro comportamento per evitare i danni potenzialmente derivanti da una cattiva reputazione. Un esempio molto noto è quello di *TurkOpticon*, piattaforma indipendente di rating creata dagli utenti di *Amazon Mechanical Turk*. Sul punto v., tra gli altri, M. Silberman, *Operating an employer reputation system: Lessons from Turkopticon, 2008-2015*, in *Comparative Labour Law & Policy Journal*, 2015, 37, p. 505 ss.

6. Conclusioni

Lo scenario descritto signora offre spunti, i margini per alcune brevi considerazioni finali. Come abbiamo visto, a dispetto dell'espressione che la definisce e dell'idea che probabilmente domina l'immaginario collettivo, la discriminazione algoritmica è il risultato di processi che trascendono l'operato delle sole macchine, il frutto di dinamiche complesse, alimentate da fattori umani e sociali oltre che tecnologici che è necessario imparare a esplorare e regolare in modi nuovi.

All'interno di questo scenario, gli algoritmi sono non solo causa del problema ma anche potenzialmente parte della soluzione e ciò per due ragioni di fondo. La prima: le nuove forme di analisi critica e regolazione in grado di contribuire a contrastare il fenomeno passano tutte, in un modo o nell'altro, attraverso gli algoritmi. Il discorso vale tanto per i metodi delle scienze sociali computazionali destinate ad alimentare la prospettiva dei *critical data studies*, quanto per le diverse declinazioni del paradigma tecnoregolativo. La seconda: proprio in quanto mediata dagli algoritmi, la discriminazione lascia tracce digitali che, debitamente raccolte e analizzate, consentono di trovare negli algoritmi la soluzione per smascherarla e contrastarla⁵².

Di fronte a questa prospettiva, le sfide da affrontare sono considerevoli non solo sul piano delle scelte *stricto sensu* normative, ma anche su quello scientifico e culturale. In un universo in cui i diritti sono condizionati dalla tecnologia e dalla sua pervasiva penetrazione in ogni aspetto della vita sociale ed economica, servono programmi di ricerca che facciano interagire in modi nuovi i tanti saperi - dal diritto all'informatica, dalla teoria della complessità alle scienze sociali - interrogati dall'obiettivo di valutare e contrastare le ingiustizie algoritmiche.

In quest'ottica, i giuristi sono chiamati ad uno sforzo di apertura al dialogo con prospettive, conoscenze e metodi appartenenti ad ambiti disciplinari eccentrici rispetto al mondo del diritto. Non solo. A loro spetta anche attrezzarsi tecnicamente per confrontarsi con un nuovo compito: imparare a tradurre le proprie esigenze applicative e conoscitive nella progettazione di "macchine", artefatti tecnologici da utilizzare e, prima ancora, da costruire in collaborazione con studiosi di altre aree di ricerca.

Il cambiamento richiederà senza dubbio tempo e un graduale processo di fertilizzazione incrociata che consenta di modificare tradizioni culturali e metodologiche consolidate. Per quanto strano possa apparire, al giurista serve l'audacia di vedersi nel ruolo insolito di progettista e inventore. Parafrasando l'immagine utilizzata in un'intervista di qualche anno fa da Kevin Kelly⁵³ per descrivere il suo rapporto con la tecnologia, la sfida da affrontare per contrastare la distopia delle macchine non è quella di immaginare utopie ma quella di costruire prototipi.

52 La prospettiva è ben rappresentata in: J. Kleinberg, J. Ludwig, S. Mullainathan, C.R. Sunstein, *Algorithms as discrimination detectors*, in *Proceedings of the National Academy of Sciences*, 2020, 117 (48), pp. 30096-30100.

53 Ex direttore della rivista *Wired* ed esponente storico della cybercultura, Kelly afferma: "I'm not a utopian. I am not a dystopian. I'm a protopian". Il testo integrale dell'intervista è disponibile al link <https://bit.ly/3Lrn72W>.