



**FORUM
DISUGUAGLIANZE
DIVERSITÀ**

MATERIALI

**15 PROPOSTE PER
LA GIUSTIZIA SOCIALE**

**Ispirate dal Programma
di Azione di Anthony Atkinson**

**ALGORITMI DI APPRENDIMENTO AUTOMATICO E UTILIZZO DI DATI PERSONALI:
UNA BIFORCAZIONE SBILANCIATA
(non paper)¹**

Le tecnologie dell'informazione non hanno un effetto predeterminato sulla giustizia sociale. Il ForumDD condivide con Anthony Atkinson questo convincimento. L'effetto dipende da come quelle tecnologie vengono impiegate e a quale "cambiamento tecnologico" esse danno vita. Un "cambiamento tecnologico" non comprende, infatti, solo le tecniche per risolvere i problemi, ma i problemi stessi che si intende risolvere (G.Dosi): ha dunque anche un contenuto politico. Anche per gli algoritmi di apprendimento automatico e per l'utilizzo massiccio di dati personali e identitari riversati in rete vale la stessa cosa. Esiste una biforcazione fra un uso che accresce e uno che riduce la giustizia sociale.

E tuttavia, in parte per l'abbrivio prodotto dalle scelte compiute in questi anni e per il loro condizionamento delle future decisioni, questa particolare biforcazione appare oggi sbilanciata a sfavore della giustizia sociale. Il vero e proprio "sistema tecnologico di decisioni automatiche" che si è affermata è segnato da una straordinaria concentrazione di controllo sulla conoscenza, da cui discende una concentrazione di potere e di ricchezza, in una spirale ricorsiva che crea disuguaglianze e minacce per la democrazia.

Eppure una strada diversa sembra possibile. La reazione degli ultimi anni, le nuove idee, le sperimentazioni, la vivacità del confronto lo suggeriscono. Ma bisogna agire presto. Bisogna mettere sul tavolo un'Agenda di cose da fare, di esperimenti da realizzare, di norme già esistenti a cui dare attuazione, di alternative da promuovere. L'Italia è assai indietro su questo terreno. Bisogna recuperare il terreno perduto.

La questione generale

Lo sviluppo delle tecnologie dell'informazione, la rapidità di elaborazione e la possibilità di utilizzare masse di dati (*big data*) forniti dagli utenti della rete hanno consentito un forte sviluppo dell'intelligenza artificiale e segnatamente degli algoritmi di apprendimento automatico (d'ora in poi algoritmi) – o *machine learning algorithms*. Il loro utilizzo tocca oggi tutte le dimensioni della nostra vita: il lavoro (reclutamento, valutazione, orari), le politiche pubbliche (assistenza sociale, sicurezza, giustizia), la politica (informazione, propaganda), e soprattutto il consumo di servizi e beni sul mercato (prodotti in rete e pubblicità, intrattenimento, servizi di trasporto e turistici, credito, assicurazioni). E' su quest'ultimo terreno che essi hanno accresciuto il nostro "benessere" e hanno così raccolto il nostro spesso incondizionato consenso, permettendoci: di conoscere all'istante le opzioni disponibili di un prodotto desiderato e poi di acquisirlo con tempestività; di offrire sul mercato prodotti e incontrare la "propria" domanda; di intrattenersi (due terzi dei bytes in rete sono destinati a intrattenimento).

Algoritmi e *big data* formano un vero e proprio "sistema tecnologico di decisioni automatiche". Gli algoritmi di apprendimento automatico sono modelli che, per un dato obiettivo, utilizzando e continuamente aggiornando grandi masse di dati (numeri, parole, immagini, etc.), ricercano correlazioni fra fenomeni e forniscono previsioni su cui basare decisioni: le e-mail da eliminare come "spam", stante le e-mail eliminate in passato; le parole da associare a suoni vocali, stante le associazioni giuste e sbagliate del passato; le persone da assumere, stante i risultati passati di "persone con simili caratteristiche"; il messaggio pubblicitario o politico da inviare a una data persona, stante le preferenze mostrate in passato da "persone con simili caratteristiche". La previsione di un algoritmo e la conseguente decisione sono determinate dall'*obiettivo* che ci si è proposti – ridurre le e-mail da visionare, minimizzare le associazioni sbagliate di parole a suoni, minimizzare le assunzioni di persone non coerenti con i risultati attesi, assecondare le preferenze dei destinatari del messaggio - e dai *dati che si sceglie di utilizzare*. Gli algoritmi non ricercano le relazioni causali fra i fenomeni, ma le correlazioni fra essi. Nel fare questo, essi incorporano i *valori* di chi li impiega, riflessi nell'obiettivo e nei dati.

¹ La nota è stata redatta da Fabrizio Barca per il Seminario tenutosi a Messina l'1 febbraio 2019 (cfr. Materiali). E' stata quindi modificata e ampliata anche sulla base dei risultati di quel confronto e al contributo inviato da Alberto Cottica.

Consentendo di tenere conto, anche in tempo reale, di uno straordinario volume di informazione e offrendo nuove e talora straordinarie opportunità di profitto a chi li impiega, gli algoritmi si sono diffusi in tutti i campi della nostra vita. Dal punto di vista dei cittadini e consumatori, essi hanno ampliato la gamma, ridotto i tempi e spesso ridotto il prezzo di servizi di cui quotidianamente molti si avvalgono: dall'ordine di prodotti via rete all'affitto di stanze, dal trasporto alla selezione delle opportunità di lavoro, fino alla ricerca di dozzine di materiali utilizzati per scrivere queste pagine. Dal punto di vista dell'impatto sulla giustizia sociale, come molte altre "tecniche", gli algoritmi non sono in sé né "giusti" né "ingiusti". Dipende da come sono utilizzati e precisamente: dal fatto che considerazioni di giustizia sociale (incluse considerazioni a favore della concorrenza sul mercato) siano incorporate nei valori e obiettivi del modello e nella scelta dei dati; dalla possibilità che chi sarà poi influenzato dall'uso delle loro previsioni possa pesare sulla scelta di quei valori e di quei dati; e, prima ancora, che egli/ella possa controllare l'uso dei dati che fornisce. Ma rispetto ad altre tecnologie sviluppate nell'ultimo decennio, *la biforcazione aperta dal sistema tecnologico di decisioni automatiche fondato su algoritmi e big data appare oggi sbilanciata a sfavore della giustizia sociale.*

Il paradosso e gli effetti sistemici degli AAA

Gli algoritmi di apprendimento automatico sono al centro di un paradosso. In molti casi essi sono stati introdotti come strumenti, non solo per accrescere l'efficienza nel prendere decisioni, ma proprio per evitare ingiustizie sociali: discriminazioni soggettive nelle assunzioni, nell'erogazione del credito o nella comminazione di pene. La motivazione era semplice: sottraendo tali decisioni a una valutazione soggettiva, a volte prona a discriminazioni (di genere, di origine etnica, di apparenza, etc.) e a favoritismi, e affidandole a "meccanismi automatici", si avranno meno iniquità.

Le cose hanno preso una diversa piega, per motivi insiti nella natura stessa degli algoritmi. Se nell'elaborare i dati del passato, le caratteristiche personali che appaiono correlate negativamente con l'obiettivo fissato (luogo di vita, persone frequentate, livello di istruzione) risulteranno concentrate proprio fra i ceti deboli, è fra essi che gli algoritmi condurranno a concentrare le non-assunzioni, il rifiuto del credito, le pene più elevate. Dal momento che gli algoritmi suggeriscono decisioni per il futuro che ricalcano il passato, in un contesto di elevate disuguaglianze essi tendono a riprodurre e rendere sistematiche quelle disuguaglianze. Ad esempio, negando sistematicamente credito ai poveri – leggi: a persone le cui caratteristiche risultano coincidere con questo gruppo sociale – dal momento che essi risultano in passato più proni a non restituire crediti, gli algoritmi accrescono la loro incapacità di onorare i crediti. Ma questa è solo una parte della storia.

La riproduzione delle disuguaglianze attraverso i meccanismi decisionali è sempre esistita. Non è questa la novità degli algoritmi di apprendimento automatico, che anzi, come visto, possono eliminare discriminazioni soggettive. La novità e i rischi per la giustizia sociale stanno nella scala e nell'incontrollabilità del meccanismo decisionale automatico (le correlazioni non hanno "motivazione" o "logica") e in alcuni altri tratti che possiamo così riassumere:

- ! *Crisma di oggettività.* L'apparente natura "tecnica" degli algoritmi e la potenza elaborativa dona loro un crisma di oggettività che ne impone le previsioni come fossero "leggi" da eseguire senza discussione, indipendenti da scelte soggettive, tendendo a precludere confronto e partecipazione².
- ! *Scala sistemica degli effetti perversi.* La scala di applicazione degli algoritmi e la loro replicabilità sono talmente forti da poter produrre effetti sistemici che convalidano ex-post le previsioni fatte: ad esempio la negazione del credito o dell'assunzione a una categoria di persone classificate come inaffidabili, peggiora le condizioni di quella categoria producendo comportamenti che confermano le previsioni. Questo fenomeno, tipico anche dei meccanismi decisionali tradizionali, può assumere una diffusione e una rapidità tali da avere effetti sistemici. Il potere che algoritmi impiegati in modo cieco a considerazioni di giustizia sociale hanno di autorealizzare le proprie previsioni è assai elevato.
- ! *Pregiudizi istintivi replicati e amplificati.* E' stato argomentato e dimostrato che gli algoritmi tendono a replicare i pregiudizi istintivi che sono racchiusi nel linguaggio che alimenta gli algoritmi

² Sulla retorica dell'automazione e dell'intelligenza artificiale e il suo ruolo ideologico, cfr. A. Taylor, *The Automation Charade*, Logic Magazine, oct 2, 2018.

stessi.³ Il nostro comportamento, le nostre opinioni e i nostri orientamenti possono assumere forme che si allontanano da tali pregiudizi istintivi; ma se le decisioni sono invece affidate a meccanismi automatici che replicano quei pregiudizi, le distorsioni in termini di giustizia sociale vengono amplificate.

- ! *“Scatola nera” non verificabile.* Gli algoritmi sono di proprietà di chi li ha costruiti e la logica delle previsioni (e “decisioni”) che producono, non essendo fondate su teorie, non sono comprensibili neppure agli utilizzatori: sono vere e proprie “scatole nere” (*black box*). Assai più che con i modelli previsivi del passato, gli utilizzatori stessi (imprese, scuole, tribunali, banche, etc.) non ne conoscono il funzionamento: essi divengono così impermeabili a richieste di chiarimento o di contestazione. A questa perdita di verificabilità corrisponde il profitto di chi è proprietario dell’algoritmi.
- ! *Disumanizzazione delle politiche e negazione del “riconoscimento”.* La natura di “scatola nera” produce conseguenze ancor più gravi quando gli algoritmi sono utilizzati per la produzione di servizi di mercato o, peggio, di servizi di cura della persona in cui la persona stessa cerca nel rapporto con il fornitore del servizio il “riconoscimento” delle proprie personali condizioni. Nel caso delle politiche per la povertà, la relazione con gli assistenti pubblici incaricati di verificare la natura e il merito dell’intervento è parte integrale del servizio: se tali funzioni sono affidate a un “robot” che classifica e giudica (o anche dialoga con) ogni persona come un anonimo membro del gruppo in cui gli algoritmi la collocano, quel riconoscimento umano, fra persone, viene meno; la natura del servizio ne risulta stravolta.
- ! *Segmentazione in micro-gruppi destinatari di messaggi dedicati.* Sia nel mercato che nell’arena politica, è da sempre invalsa la pratica di segmentare il pubblico in gruppi in relazione alle loro supposte preferenze, per renderli destinatari di messaggi (pubblicitari o politici) dedicati. L’utilizzo degli algoritmi applicati a grandi masse di dati consente di rendere estremamente granulare la segmentazione. Abbinata al rapporto 1 a 1 fra destinatario e produttore del messaggio, consente di isolare ogni destinatario o gruppo di destinatari dagli altri, scoraggiando un confronto di messaggi, la loro contestabilità e il confronto pubblico e aperto su di essi.
- ! *Uso non retribuito dei dati identitari immessi dagli utenti nella rete.* Della massa di dati utilizzati dagli algoritmi fa parte la mole di informazioni sulla nostra identità che ogni giorno riversiamo, spesso inconsapevolmente, in rete. Essi sono impiegati non solo dai giganti che controllano gran parte del *social network* – Google e Facebook, in testa – ma anche da tutte le imprese che li acquisiscono direttamente dagli utenti della rete o da altre fonti per poi impiegarli negli algoritmi. Si compie in questo modo uno “scambio ineguale”, o comunque né negoziato né regolato, fra utenti e imprese, che concorre ad accumulare in poche mani enormi profitti. Ma non basta. L’ingiustizia sociale si manifesta nell’assenza di qualunque forma di controllo dei produttori di dati personali sul modo in cui essi saranno impiegati.

Siamo in conclusione in presenza di profonde ragioni fisiologiche per cui, anche trascurando ogni patologia, l’uso degli algoritmi può produrre gravi effetti negativi sulla giustizia sociale. E veniamo all’uso dei dati identitari, che degli algoritmi sono l’alimento.

Il sistema di produzione e utilizzo dei dati identitari in rete

Il tema della produzione, dell’utilizzo e del controllo dei dati identitari – dati connessi alla nostra cittadinanza o a servizi pubblici di cui ci avvaliamo, altri dati identificativi, opinioni, immagini, contatti, itinerari, etc - che riversiamo in rete rappresenta, come noto, uno degli aspetti più delicati e controversi dell’impatto dell’attuale cambiamento tecnologico sulla giustizia sociale, strettamente legato allo sviluppo dell’intelligenza artificiale.

La tecnologia dell’informazione modifica radicalmente le modalità della nostra “identificazione” nella partecipazione ai doveri e diritti di cittadinanza, ai mercati (lavoro, consumo, credito) e a ogni sistema complesso di relazioni. Da un lato, essa rende più facile e ha incoraggiato l’attribuzione di “identità” a milioni di essere umani che non l’avevano e che di conseguenza non avevano diritti⁴. E, come si è ricordato,

³ Cfr. Caliskan-Islam, J.J. Bryson, A. Narayanan, *Semantics derived automatically from language corpora necessarily contain human biases*, mimeo, Princeton University-University of Bath, August, 2018. Ci ritorniamo più Avanti nel testo.

ha consentito lo sviluppo di nuovi servizi, tempestivi e mirati sulle esigenze personali. Dall'altro, ha aperto problemi nuovi per la giustizia sociale:

- ! l'accresciuta possibilità di contraffazione di questi dati;
- ! l'opacità e non verificabilità del loro uso da parte di "centri" di gestione, privati o pubblici;
- ! Le violazioni della privacy e i rischi personali derivanti dalla rilevazione dei dati relativi ai nostri spostamenti georeferenziati raccolti attraverso App - con scarsa possibilità di evitarlo e l'assenza di protezioni formali sull'uso;⁵
- ! la difficoltà o l'impossibilità di accedere al complesso di questi dati – i "centri" possono farlo, non le persone che li forniscono;
- ! la difficoltà o l'impossibilità di sanzionare contraffazioni, errori, utilizzi o cessione a terzi dei dati, anche se espressamente vietati⁶;
- ! l'elevato standard dell'onere di prova, e il costo e la complessità necessari per provare un "errore" che risulti contenuto nei dati identitari utilizzati da un algoritmo e che abbia condotto a una decisione avversa - nell'ottenere un credito o nella selezione per un lavoro – rispetto allo standard richiesto allo stesso algoritmo per le sue previsioni;
- ! i rischi, solo in parte già apprezzabili, di essere – noi utenti della rete - suddivisi in micro-gruppi target di cui vengono monitorati i comportamenti e che ricevono messaggi di mercato o politici differenziati, inconsapevoli gli uni degli altri⁷.

Il tema della proprietà e del controllo dei dati identitari utilizzati dagli algoritmi è dunque centrale.

Utilizzando dati in larga misura personali, i giganti della rete e altre imprese che acquisiscono tali dati realizzano straordinari profitti e un forte ed esclusivo potere di controllo, di ricerca e di sviluppo di nuove applicazioni. Il vantaggio acquisito nell'accesso al patrimonio comune di dati identitari concede loro una posizione monopolistica difficilmente scalfibile. E qui nasce lo scambio ineguale con gli utenti. Nei loro confronti le imprese possono così agire, come scrivono due esponenti del pensiero liberale radicale, Eric Posner e Glen Weyl, come "tecno-feudatari": come "gli aristocratici si impossessavano della parte di qualità del prodotto agricolo dei propri servi", "lasciando loro abbastanza per sopravvivere", le "sirene dei server – come le definiscono – offrono servizi utili e godibili, impossessandosi del valore di mercato che noi in cambio produciamo".⁸ E c'è altro. Il sistema così costruito, per verificare, decodificare e completare le informazioni fornite, integra il lavoro non remunerato degli utenti con lavoro mal pagato di grandi masse di precari (*crowd workers*), la cui attività viene tenuta in un cono d'ombra, per non rovinare l'immagine pulita e magica dell'"intelligenza artificiale".⁹

La sovranità digitale è insomma trasferita ad alcuni grandi monopoli, configurando un vero e proprio modello di "sovranità dei monopoli" che ha negli USA il suo punto più avanzato. L'alternativa di una "sovranità dello Stato" che viene dalla Cina appare altrettanto pericolosa. Il regime è di nuovo di assoluta centralizzazione, ma in questo caso è l'apparato dello Stato a governare dati e algoritmi. Il piano 2014-2020 per costruire il Sistema del Credito Sociale prevede di utilizzare dati e algoritmi per classificare la reputazione dei cittadini e delle imprese, in tutti i campi. La sua sperimentazione ha già permesso al governo cinese di negare nel 2018 a oltre 17 milioni di cittadini cinesi l'acquisto di un viaggio aereo e a oltre 5 milioni l'acquisto di un viaggio in treni ad alta velocità come sanzione per violazioni, come la falsa pubblicità, l'abuso di droghe o il mancato pagamento di multe¹⁰.

4 Come osserva *AI-Now Report 2018* – AI Now è un istituto di ricerca USA che studia gli effetti sociali dell'intelligenza artificiale - l'importanza di questo obiettivo per la giustizia sociale è rimarcato dal fatto che esso è previsto espressamente dai *Sustainable Development Goals* dell'ONU (obiettivo 16.9).

5 Cfr. ad esempio: <https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html>.

6 Per il peso crescente di questo aspetto, esploso nel 2018, cfr. *AI-Now Report 2018*.

7 Su questo punto cfr. Il caso estremo riportato in C. O'Neil, *Weapons of Math Destruction*, Broadway Books, 2016, pp. 171-173.

8 Cfr. E.A.Posner, E.G. Weyl, *Radical Markets*, Princeton University Press, 2018, p. 231.

9 Cfr. M.L.Gray, S. Suri, *The Humans Working Behind the Curtains*, Harvard Business Review, jan 2017 e A. Taylor (2018). Sulle condizioni di questi lavoratori e lavoratrici, cfr. AI Now – Report (2018).

Per tutte queste ragioni, è necessario intervenire. Il “sistema tecnologico di decisioni automatiche” costituito dagli algoritmi e dall’uso dei nostri dati identitari può produrre profondi e crescenti effetti negativi sulla giustizia sociale, concorrendo a concentrare grandi poteri in poche mani e a polarizzare la società. È utile richiamare brevemente gli effetti che possono prodursi in quattro distinte dimensioni della nostra vita

Per tutte queste ragioni, è necessario provvedere. Gli algoritmi e i nostri dati identitari formano oggi un “sistema tecnologico incontrollato” che può produrre profondi e crescenti effetti negativi sulla giustizia sociale, concorrendo a polarizzare la società fra vincitori e perdenti. La crescita di questo sistema tecnologico ha un ritmo esponenziale, toccando tutti i campi della nostra vita. La consapevolezza di questo stato di cose è cresciuta, come mostrano le ricerche a cui questa nota ha potuto fare riferimento. Il confronto è cresciuto, così come le idee volte ad affrontare la questione e la sperimentazione dei rimedi, specie nei paesi che sono sulla frontiera dell’uso di questi modelli. E’ necessario che ciò avvenga anche in Italia.

Prima di valutare strade e ipotesi diverse di intervento che si vanno configurando, è utile richiamare, con alcuni esempi tratti dal dibattito internazionale in atto, le principali modalità con cui il problema si manifesta oggi in quattro distinte dimensioni della nostra vita.

I rischi per la giustizia sociale in tutte le dimensioni di vita

Lavoro

La raccolta di dati personali e il loro utilizzo da parte di algoritmi di apprendimento automatico si stanno diffondendo in modo massiccio in tutti i profili del lavoro e del non-lavoro: nelle diverse fasi della selezione del personale, nella sua valutazione (a fini di carriera o di licenziamento), nella determinazione dell’orario di lavoro; ma anche nel modo in cui utilizziamo il nostro tempo di vita, consumando, viaggiando, intrattenendoci, imparando, prendendoci cura di altro e altri. Trattiamo di questo secondo aspetto più avanti, parlando di consumo (visto che larga parte di queste attività avvengono attraverso il mercato)

In forte diffusione da tempo è il ricorso ad algoritmi nella prima fase dei processi di *selezione del personale*, per restringere la lista di coloro che verranno sottoposti a più attenta verifica. Dal punto di vista della giustizia sociale, tale ricorso presenta il vantaggio di superare i giudizi soggettivi discriminatori o influenzati da relazioni. Ma a fronte di ciò sta il rischio di reintrodurre *de facto*, come si è visto in generale, tali discriminazioni per via delle caratteristiche degli offerenti che vengono considerate nelle elaborazioni: prescindendo dai tratti unici di ogni singolo candidato, questo viene attribuito a un dato “gruppo”, ritenuto più o meno adatto alle esigenze dell’impresa. Questi rischi si accentuano in particolare con il ricorso, fra i dati impiegati, agli esiti di test sulla personalità¹¹ o di colloqui condotti da “simulatori di conversazione” (*chat bot*), come avviene già anche in Italia.

Nella fase successiva delle selezioni, il ricorso ad algoritmi nella lettura e valutazione dei Curriculum Vitae o di altra documentazione può esasperare la tendenza ad affidarne la redazione a soggetti specializzati in grado di anticipare i criteri di raggruppamento automatico degli algoritmi: ne risultano favoriti i candidati che dispongono dei mezzi finanziari e delle relazioni per avvalersi di tali servizi. Più in generale, la tendenza a impiegare questi metodi solo per le posizioni con minori competenze (o comunque con maggiore offerta) crea un divario rispetto alle posizioni più elevate, per le quali permane il ricorso a metodi basati su relazioni umane.

L’utilizzo degli algoritmi nella *valutazione del personale* incorre in rischi connessi ai dati utilizzati. Se, ad esempio, si valuta la capacità innovativa sulla base delle “nuove idee” rintracciate in forme codificate (note o proposte registrate in momenti collettivi) si rischia di penalizzare soggetti magari più innovativi ma che manifestano le loro idee in modo “anomalo” o cooperativo, condividendole informalmente con gli altri. L’”ingiustizia” non riguarda in questi casi una discriminazione sociale o di genere, ma il fatto che a soggetti

¹⁰ Cfr. *South China Morning Post*, 19 febbraio 2019 (<https://www.scmp.com/economy/china-economy/article/2186606/chinas-social-credit-system-shows-its-teeth-banning-millions>).

¹¹ Cfr. O’Neil (2016), p.108.

con alcuni tratti comportamentali non viene data la stessa opportunità che ad altri. Decisamente allarmante è, poi, l'utilizzo che può essere fatto di informazioni raccolte in tempo reale attraverso strumenti di sorveglianza. Si tratta di strumenti che iniziano a essere impiegati¹² e che consentono di registrare, per ogni dipendente, conversazioni, movimenti, posture e prossimità ad altri, combinandoli con i messaggi su rete effettuati da ogni singolo lavoratore: alle minacce per la privacy, si aggiunge la possibilità di costruire sulla base di questi dati metodi o pressioni coercitive sul lavoro.

Già assai manifeste e gravi sono, infine, le conseguenze che tende ad avere l'applicazione degli algoritmi nella determinazione degli *orari di lavoro*. La disponibilità di dati sulla dinamica della domanda in un'attività di servizio rivolta al pubblico e di tutte le variabili che si presume la influenzino e la loro elaborazione continua attraverso algoritmi rendono possibile prevedere l'offerta di lavoro necessaria momento per momento durante ogni giornata e aggiornare continuamente tale previsione. Ne possono derivare e in molti contesti del comparto terziario ne sono già derivati, non solo l'intensificazione dei ritmi di lavoro – non ci sono più momenti “tranquilli” – ma soprattutto la pratica di comunicare ai/alle dipendenti con brevissimo periodo di preavviso le necessità di presenza, rendendo per essi/esse impossibile la programmazione dei tempi di vita. E' evidente che questo *just in time* del lavoro è dettato dall'aver assegnato agli algoritmi utilizzati il solo obiettivo del massimo profitto, senza alcuna attenzione per la qualità di vita (e di lavoro) dei lavoratori: il problema non sta, quindi, ancora una volta, nello strumento ma negli obiettivi a cui viene piegato. E' un impiego che colpisce ovviamente solo i lavoratori con minori competenze, addetti a questi lavori, e le loro famiglie.¹³

Servizi pubblici essenziali e politiche di assistenza sociale

In questo ambito i segnali di rischio venivano finora soprattutto dagli USA¹⁴. Ma i primi utilizzi di un software di riconoscimento facciale denominato Sistema automatico di un riconoscimento immagini (SARI), costruito in un'ottica anti-terrorismo e in dotazione alla polizia italiana e la raccolta massiccia di informazioni identitarie per il nuovo cosiddetto reddito di cittadinanza con finalità di sorveglianza per evitare abusi nell'uso dello strumento, fanno intravedere un crescente impiego degli algoritmi nell'azione pubblica. E' dunque utile avere ben presenti le lesioni per la giustizia sociale indotte negli USA dal ricorso agli algoritmi nelle politiche del welfare. Consideriamo in particolare gli ambiti della giustizia, della sicurezza e dell'assistenza sociale (a persone con disabilità o in condizioni di povertà).

Nel caso della *giustizia*, è noto l'impiego in molti Stati USA degli algoritmi per prevedere la probabilità di recidiva di un soggetto che ha commesso un reato, legando a questa previsione la misura della sentenza. Si esasperano qui i rischi di discriminazione (derivanti dalla correlazione delle variabili considerate con l'origine sociale o etnica delle persone) e di autorealizzazione delle previsioni. Tale utilizzo è impedito in Italia sia dal Codice penale (art. 220, comma 2), sia dal Regolamento Europeo sulla protezione dei dati personali, secondo cui nessun atto che implichi la valutazione di comportamenti può essere fondato unicamente sul trattamento automatico di dati personali volti a definire profilo e personalità della persona (art.22).

Per quanto riguarda la *sicurezza*, è diffuso, sempre negli USA, l'impiego degli algoritmi per identificare le zone di una città dove concentrare il contrasto della criminalità. A utilizzi efficaci e di successo, che non toccano profili di giustizia sociale, si accompagnano degenerazioni quando fra i dati sono stati inseriti illegalità minori (conducendo a esacerbare tensioni sociali), o quando l'identificazione delle aree ha condotto a vere e proprie azioni preventive intimidatorie. Quest'ultima opzione rischia di essere accentuata gravemente dal ricorso ad algoritmi per il riconoscimento facciale.¹⁵ In Italia, come si accennava,

¹² Si veda in proposito *The Economist*, in un breve servizio (<https://www.facebook.com/TheEconomist/videos/602515400213741/>) sugli strumenti di sorveglianza prodotti dalla società paradossalmente denominata “Humanize”.

¹³ Cfr. O'Neil (2016), p. 129, anche per gli effetti sui figli di lavoratori in questa situazione.

¹⁴ Cfr. V. Eubanks, *Automating Inequality*, St. Martin's Press, 2018.

¹⁵ Cfr. *AI-Now Report 2018*, pp. 12-17.

nell'ottobre 2018, ha avuto luogo un arresto grazie al fatto che il suo volto ripreso da una telecamera è stato associato al suo nome dalla polizia grazie a un software di riconoscimento facciale denominato Sistema automatico di un riconoscimento immagini (SARI), costruito in un'ottica anti-terrorismo e in dotazione alla polizia: un'interrogazione parlamentare (alla quale non risulta sia stata data risposta, <https://parlamento18.openpolis.it/atto/documento/id/177337/>) ha chiesto chiarimenti sulla dimensione dei soggetti schedati, sull'impiego, sulla sua compatibilità con la Regolamentazione Europea (di cui oltre).

La diffusione nell'uso dei sensori collocati nelle auto e nelle strade per monitorare il traffico e le modalità di guida può avere effetti positivi sulla mobilità e sugli incidenti stradali. Ma se anche questa massa di dati in tempo reale dovessero essere appropriata da poche imprese digitali, esse acquisirebbero un potere che si presta a ogni possibile utilizzo e che tende a trasferire decisioni pubbliche in mani private.

Per quanto riguarda *l'assistenza sociale*, il caso USA mostra come, sotto la pressione dei tagli di bilancio, gli algoritmi possono essere impiegati per tagliare trasferimenti alle persone e stravolgendo la natura stessa del servizio di assistenza. E' il caso di riduzioni sistematiche del numero di ore di assistenza a persone con disabilità, che avviene senza darne le motivazioni, motivazioni che restano racchiuse nella "scatola nera" di un algoritmo¹⁶. Più in generale, l'analisi sistematica dell'impatto dell'uso di algoritmi nella gestione dell'assistenza ai poveri, a persone senza abitazione o in condizioni di indigenza¹⁷, mostra che, sempre con l'obiettivo di realizzare tagli delle spese, gli algoritmi hanno finito per assecondare un atteggiamento punitivo nei confronti delle persone povere.

Ad una sistematica violazione della privacy, si associa il venire meno di un rapporto personale fra assistente e assistito e dunque di un'attenzione umana in grado di valutare i singoli casi e di segnalare empatia e riconoscimento delle condizioni della persona: se questo, come per altre applicazioni, elimina il rischio che l'erogazione del servizio sia soggetta a favoritismi, elimina anche e più gravemente una componente fondamentale del servizio stesso (il riconoscimento e la relazione umana) e produce mortificazione sistematica. Ogni persona è ben consapevole che non verrà considerata per ciò che è, per la propria unicità, ma perché, in base alle proprie caratteristiche, essa risulta appartenere a un dato micro-gruppo sociale, che merita o non merita l'intervento.

Consumo di servizi e beni sul mercato

Qui siamo nel campo di massimo utilizzo degli algoritmi di apprendimento automatico. Una larga parte del nostro tempo di vita non dedicato al lavoro avviene "attraverso il mercato": scegliendo, acquistando, consumando beni e servizi, viaggiando, intrattenendoci, imparando, prendendoci cura di altro e altri. E' questo il campo dove l'impiego del "sistema tecnologico delle decisioni automatiche" ha la sua massima e più manifesta applicazione. E' su questo terreno che esso ha accresciuto il nostro "benessere" e ha dunque comprato il nostro consenso, permettendoci di: conoscere all'istante le opzioni disponibili di un prodotto desiderato e poi di acquisirlo con tempestività; offrire sul mercato prodotti e incontrare la "propria" domanda; intrattenersi (due terzi dei bytes in rete sono destinati a intrattenimento).

Ma questo è anche il campo dove si manifestano molteplici gravi rischi:

- ! In primo luogo, si manifesta qui in modo eclatante lo scambio ineguale e comunque non negoziato, di cui si è detto, fra utenti della rete che forniscono informazioni e imprese che in cambio erogano servizi.
- ! In secondo luogo, è il campo dove il ricorso agli algoritmi consente un forte condizionamento delle nostre preferenze: le imprese che controllano i principali canali informativi della rete e in genere tutte quelle che offrono prodotti in rete utilizzano le informazioni derivanti dalle nostre decisioni o che ci vengono all'uopo richieste per profilare le nostre caratteristiche e suddividerci in modo granulare in gruppi che saranno poi destinatari inconsapevoli di messaggi differenziati volti a influenzare le preferenze.

¹⁶ Cfr. *AI-Now Report 2018*, pp. 18-19.

¹⁷ Si vedano le cronache e analisi di V. Eubanks, *Automating Inequality*, St. Martin's Press, 2017.

- ! In terzo luogo, nel campo della cultura e dell'intrattenimento, il processo di fortissima concentrazione dell'industria creativa, nel determinare un condizionamento sistematico delle preferenze (che si legano fortemente ai rischi sotto richiamati sul terreno dell'informazione della politica), può ridurre la diversità culturale e condizionare artisti, creativi e in genere l'offerta culturale.
- ! In quarto luogo, il ricorso ad algoritmi per la determinazione dei prezzi può favorire una collusione oligopolistica fra le imprese a danno dei consumatori.
- ! Infine, nel caso di alcuni servizi di mercato costruiti su relazioni fiduciarie, segnatamente credito e assicurazioni, il ricorso agli algoritmi può condurre a discriminazioni sistematiche (fino alla negazione *de facto* del servizio) e ad un vero e proprio stravolgimento della natura del servizio stesso.

Gli ultimi tre rischi meritano particolare attenzione.

Anche nell'industria creativa, il processo di concentrazione della produzione attraverso piattaforme (cresciute, come sempre, prima disintermediando e poi re-intermediando il mercato) è stato potente: prima di tutto Netflix e poi Amazon, Disney, Comcast, Baidu, Tencent, Youku/Alibaba dominano il mercato. Gli algoritmi servono a queste mega-imprese soprattutto per ricercare, selezionare e governare i veri produttori (gli artisti, i creativi) e per analizzare il mercato e distribuire i prodotti che sembrano maggiormente soddisfarne (o di cui sono in grado di orientarne) i gusti (discovery). La scala raggiunta da queste piattaforme dà loro una forte e difficilmente scalabile posizione oligopolistica. Come scrivono XXX, da cui sono tratte queste considerazioni: “Quando decide di prolungare una serie TV, Netflix elabora i dati pregressi: il numero e la durata delle visioni per ciascun paese e cluster di spettatori, l'incidenza sugli abbonamenti, le performance di ciascun episodio a seconda del regista, attori, sceneggiatori, le scene più appassionati e quelle più fiacche, le reazioni al sonoro e ai colori. Conosce i costi e le performance del produttore e dei produttori concorrenti che in tutto il mondo possono sostituirlo. È in grado di prevedere con discreta approssimazione i risultati della prossima stagione, in relazione ai costi e alle scelte produttive e artistiche. La controparte negoziale non ha nessuna di queste informazioni.”

Anche senza immaginare scenari distopici nell'uso di questo potere (combinato con le finalità politiche di cui oltre), ne risulta un forte potere dei pochi che controllano le piattaforme nel fissare prezzi, ripartire ritorni e stabilire i contenuti.

Ma gli *effetti restrittivi sulla concorrenza* e dunque sui prezzi riguardano tutti i mercati. La possibilità che gli algoritmi conducano a una sistematica violazione della concorrenza è tema su cui è di recente maturata l'attenzione delle Autorità preposte alla tutela della concorrenza. Un recente studio ha dato corpo a queste preoccupazioni¹⁸. Gli algoritmi per la definizione dei prezzi di vendita hanno un semplice obiettivo: massimizzare i profitti; sulla base dei dati progressivamente raccolti, gli algoritmi, senza che venga loro data alcuna istruzione, simulano strategie e sulla base delle reazioni degli altri soggetti offerenti e dei consumatori individuano la strategia di prezzo ottimale. La simulazione delle decisioni di due algoritmi attivi, per due concorrenti, sullo stesso mercato realizzata dai ricercatori ha mostrato che i due algoritmi imparano a colludere nella fissazione del prezzo: essi apprendono l'uno dall'altro, imparano a non lanciarsi in “guerre di prezzo” e si accomodano su un prezzo di oligopolio, superiore a quello concorrenziale, anche se inferiore a quello monopolistico. La gravità di questo risultato sta nel fatto che la collusione avviene in assenza di qualunque interazione fra i due algoritmi (se non attraverso i dati forniti dal mercato), senza alcuna istruzione a realizzarla e dunque senza alcuna traccia normalmente perseguita dalle Autorità che tutelano la concorrenza.

Nello specifico caso del credito e delle assicurazioni, il ricorso agli algoritmi, oltre a creare rischi per la concorrenza e la stabilità finanziaria, può produrre una trasformazione di natura strutturale del servizio.

¹⁸ Cfr. E. Calvano, G. Calzolari, V. Denicolò, S. Pastorello, *Artificial Intelligence, algorithmic pricing and collusion*, CEPR Discussion Paper, 13405, 2018.

Nel campo generale dei servizi finanziari, il ricorso agli algoritmi è cresciuto progressivamente fino a esplodere poi a partire dal 2012-13 - anche in Europa, specie in Gran Bretagna e Germania¹⁹ - con l'obiettivo primario di ridurre i costi unitari e accrescere l'efficienza. Combinati con il ricorso massiccio a dati identitari di ogni fonte, gli algoritmi trovano applicazione: nella identificazione dei segnali di più elevati rendimenti; nella valutazione del merito di credito; nella fissazione dei premi assicurativi; nella relazione con i clienti; nella sorveglianza, valutazione qualità dei dati e nell'identificazione delle frodi. Il loro impiego ha consentito la rapida crescita di un settore cosiddetto FinTech (e InsurTech) di imprese che competono con le imprese finanziarie tradizionali costruendo un ponte fra offerta e domanda di risorse finanziarie attraverso piattaforme digitali senza la necessità di immobilizzi finanziari (*thin layer fin-firms*)²⁰. I regolatori avvertono crescenti preoccupazioni soprattutto per la stabilità finanziaria e perché le nuove imprese ricadono ai limiti o fuori della possibilità di intervento dei regolatori stessi.

Anche prescindendo dai rischi per la stabilità, consideriamo questa tendenza dal punto di vista dei consumatori/utenti dei servizi. In termini medi, cioè con riferimento a un utente finale medio, la valutazione degli esperti del settore è che non esiste alcuna garanzia che alcun beneficio derivante da questo aumento di efficienza arrivi all'utenza²¹. Si confermerebbe così la tendenza stimata in un arco di 130 anni, per cui il costo unitario dei servizi per l'utenza non sarebbe mai cambiato: i benefici degli aumenti di efficienza sono stati interamente appropriati dalle imprese finanziarie²². Sono invece chiari i rischi in termini di diverso trattamento degli utenti stessi, e di giustizia sociale in genere. Consideriamoli separatamente per i due servizi di massimo interesse: credito e assicurazioni.

Nel caso del *credito*, il ricorso agli algoritmi nasce originariamente negli USA, come per altri impieghi, anche per evitare discriminazioni soggettive nella concessione del servizio. E in molti contesti il loro impiego, con l'utilizzo di dati di svariata origine, come la frequenza e regolarità dei pagamenti, ha consentito di valutare il comportamento dei clienti e il loro merito di credito quando essi erano privi di qualunque precedente record e un tempo non avrebbero quindi ricevuto un affidamento²³. Ma c'è il rovescio della medaglia. Ci riferiamo, in primo luogo, agli effetti già discussi per cui alle discriminazioni soggettive si sostituiscono discriminazioni sistematiche e al fatto che gli utenti, per via dell'effetto "scatola nera" degli algoritmi, non hanno modo di verificare e contestare le condizioni contrattuali offerte. Vi è poi l'elevato rischio di una politica di prezzo mirata su ogni singolo debitore. "Applicare gli algoritmi consente di valutare ... la disponibilità a pagare dei clienti"²⁴ e altri aspetti del loro comportamento: diventa così possibile chiedere tassi più elevati ai prenditori più affidabili e meno elevati ai meno affidabili²⁵. E' questa una delle strade con cui i benefici della maggiore efficienza vengono appropriati dalle imprese.

Ma c'è altro. I rischi toccano la natura stessa del servizio del credito. Il crescente ricorso a informazioni identitarie di ogni natura (anche provenienti dalla rete) produce una crescente distorsione: quella di interrogarsi non sul merito di credito del richiedente, ma sul merito di credito del particolare gruppo al quale le correlazioni statistiche lo associano. Questa è la caratteristica generale degli algoritmi, ma nel valutare il credito essa produce effetti particolari: essa infatti allontana definitivamente il servizio del credito dalla natura personale della relazione affidato-banca, dalla costruzione di un rapporto che non solo consenta di meglio prevedere i comportamenti dei richiedenti, ma che possa modificarli. E' questa frattura che può negare opportunità e accentuare disuguaglianze.

19 Cfr. FSB, *Artificial intelligence and machine learning in financial services*, November, 2017.

20 M. Bofondi e G. Gobbi (*The big promise of FinTech*, in *European Economy*, n.2, 2017), citano la frase significativa di un manager del settore: "Uber, the world's largest taxi company, owns no vehicles. Facebook, the world's most popular media owner, crates no content. Alibaba, the most valuable retailer, has no inventory. And Airbnb, the world's largest accommodation provider, owns no real estate."

21 Cfr. ancora M. Bofondi, G. Gobbi (2017)

22 Cfr. T. Philippon, *The FinTech opportunity*, NBER Working Papers, n. 22476, 2016.

23 Cfr. FSB (2017), p.12.

24 *Idem*.

25 Cfr. M. Bofondi, G. Gobbi (2017), p. 112.

Anche nel caso dei *servizi* assicurativi, l'uso degli algoritmi, a parte la riduzione dei costi per la valutazione o per l'assistenza ai clienti, può consentire benefici: identificando in anticipo le condizioni che possono facilitare incidenti (in ambito industriale, nelle abitazioni, nell'uso dei mezzi di trasporto, etc.), scoraggiando comportamenti che favoriscono incidenti o comunque le condizioni oggetto di assicurazione e, come nel caso del credito, consentendo di ridurre i premi assicurativi a soggetti senza record ma identificati come “a basso rischio” sulla base di una moltitudine di altri dati. Ma anche qui i rischi per la giustizia sociale sono significativi²⁶.

La raccolta di informazioni personali a cui viene subordinato il contratto può essere fortemente lesiva della privacy²⁷. La determinazione di contratti disegnati sulla persona (*customization*) “può determinare una minore comparabilità fra assicuratori, limitando così le possibilità di scelta”²⁸. Ma soprattutto l'applicazione degli algoritmi alla determinazione dei *premi assicurativi* può essere lesiva di una finalità sociale insita nello strumento assicurativo dalla sua nascita: la logica del bilanciamento dei rischi, per cui alla diversa rischiosità dello stesso evento assicurato per diverse persone corrisponde un premio assicurativo simile, calcolato sulla media degli assicurati.

Si tratta di un principio che mira a riequilibrare le circostanze diverse di vita, affinché i meno fortunati nella roulette della vita vedano in parte compensata questa circostanza da parte dei più fortunati. E' evidente che, quando il rischio dipende (come nella guida di un'auto o nella salute) anche dall'impegno delle persone a evitare rischi, questo principio andrà corretto – e da sempre è stato corretto - fino a condurre a premi differenziati, legati. Ma il ricorso agli algoritmi alimentati con masse di dati biometrici dell'utente può condurre allo stravolgimento di quell'originario principio, facendo pesare su ogni persona la diversa speranza di vita misurata su quei dati. Scrive a riguardo l'OCSE (p.27): “Sarebbe importante che le imprese di assicurazione distinguessero fra i casi in cui un cattivo stile di vita provoca cattiva salute e quelli in cui una persona nasce con problemi di salute che non possono essere in alcun modo affrontati dallo stile di vita”. Ma le cose sono più complesse di così. Perché la distinzione non è facile. E soprattutto perché il “cattivo stile di vita” può essere il risultato di circostanze sociali che conducono ad accettare lavori usuranti o che producono maggiori rischi. E' una questione non nuova, ma, come in molti altri casi, la potenza, sistematicità e non-verificabilità degli algoritmi la esaspera.

Informazione e opinioni politiche

Come nella formazione delle preferenze di consumo, anche nella formazione delle opinioni politiche (o sulle politiche) gli algoritmi di apprendimento automatico hanno radicalmente modificato lo stato delle cose, avviando un processo di trasformazione della comunicazione, della cultura e della politica di cui non è anticipabile l'esito. Gli algoritmi hanno accresciuto in modo straordinario la capacità di accumulare e reperire informazione – Google Search, il più grande motore di ricerca sul web, indicizza e rende accessibili, in relazione alle parole cercate, centinaia di miliardi di pagine web in 150 linguaggi. Mai prima nella storia è esistita una tale capacità di estrarre informazioni e di renderle disponibili al pubblico confronto. Ma molteplici sono gli usi che costituiscono un rischio per la democrazia e quindi per la giustizia sociale. Abbiamo già indicato i rischi connessi ad un abuso degli algoritmi per sorvegliare e intimidire le persone. Ci sono altri rischi.

Il rischio più eclatante consiste in un utilizzo degenerato e illecito dei dati personali identitari da noi forniti. Il caso che ha coinvolto Cambridge Analytica e Facebook è quello della violazione da parte del gestore di un'applicazione operante su Facebook (*thisisyourdigitallyfe*) dell'impegno a non trasferire la massa di informazioni raccolte: che sono poi state utilizzate da Cambridge Analytica per influenzare le campagne

26 Cfr. OECD, *Technology and innovation in the insurance sector*, 2017.

27 “Face recognition technology is used to predict factors such as chronological age, gender, smoking habits and body mass index (BMI). Based on this data, and accompanied by an activity sensor, such as a FitBit or physical activity tracker on a mobile phone, your expected life expectancy is provided. A term life offer is made based on this, and the term period can be selected by the policyholder.” (OECD(2017), p. 26. Cfr. anche <https://www.forbes.com/sites/parmyolson/2014/06/19/wearable-tech-health-insurance/#47c0083418bd>.

28 Cfr. IAIS, *FinTech Developments in the insurance industry*, feb. 2017.

elettorali britannica e statunitense. Ma abusi ancora più gravi potrebbero essere commessi se i giganti della rete o altre società che entrino in possesso dei nostri dati identitari, scegliessero di selezionare in modo politicamente distorto ma a noi ignoto le notizie che leggiamo, le informazioni che troviamo, i video che vediamo. Questa possibilità rappresenta una minaccia grave, accentuata dal fatto che nulla garantisce, qualora si manifestasse, che verremmo a saperlo e comunque entro quanto tempo²⁹.

Ma accantoniamo pure queste degenerazioni patologiche. Nella fisiologia del sistema sono racchiusi tre effetti negativi sistematici per la democrazia. Il primo effetto negativo di natura sistematica riguarda la moltiplicazione della “falsa informazione” che, anche in assenza dell’azione di vere e proprie organizzazioni eversive, può diventare il frutto del sistema di incentivi che muove la selezione dell’informazione sulla rete. Vediamo.

Scopo dei servizi offerti su rete è di massimizzare il tempo trascorso dagli utenti nell’uso del servizio così da massimizzare la pubblicità. Come osserva il recente Rapporto del Parlamento britannico *Disinformation and Fake News*, “gli algoritmi daranno sempre priorità alle storie negative dal momento che sono condivise con più frequenza rispetto alle storie positive.” Fin qui si tratta dell’amplificazione – tipica dell’uso degli algoritmi – di tendenze del passato. La novità sta nel fatto che la distorsione a favore della “frequenza di condivisione” si estende alle false informazioni. Si pensi a YouTube e al suo algoritmo (segreto) che seleziona i video raccomandati all’utente che accede al servizio. Le sue raccomandazioni saranno influenzate dagli utilizzatori compulsivi; se tali utilizzatori sono ossessionati dal fatto che la terra è piatta e compulsivamente guardano video che diffondono questa informazione, essi verranno “copiati” dall’algoritmo che diffonderà in maniera massiccia quelle stesse raccomandazioni (perché “sa” che hanno successo). E’ esattamente ciò che avviene: infatti i video sulla terra piatta risultano promossi dieci volte di più di quelli sulla terra rotonda³⁰.

In prospettiva, la produzione sistematica di falsa informazione sarà ulteriormente accresciuta dallo sviluppo dei *synthetic media*, ossia dalla diffusione di informazioni (testi, immagini, video) create o modificate da algoritmi e non distinguibili da quelle create da umani. L’exasperazione della frammentazione dei media verso cui spingerà questa forma di comunicazione potrebbe minare in modo ancora più profondo la “fiducia” nelle informazioni che ci arrivano. L’evoluzione possibile di questo stato di cose è oggetto di dibattito. Si può immaginare uno scenario di caos e di scetticismo assoluti; ovvero uno scenario di ritorno a centri o filiere di legittimazione delle informazioni, il resto delle informazioni essendo relegato a intrattenimento; ovvero, ancora, uno sviluppo di sistemi che, per ogni “fatto” o “luogo”, assemblino l’intero patrimonio informativo esistente costruendo una sorta di standard di riferimento di cosa sia “vero”; o infine, all’opposto, a forme decentralizzate di verifica e di ricostruzione di fiducia *all’interno* di singole comunità in rete (usando *blockchain*), con problemi di costruzione di fiducia *fra le* comunità³¹.

Il secondo effetto perverso di natura sistematica riguarda la possibilità, in uso da almeno un decennio, di usare algoritmi per segmentare il potenziale elettorato in gruppi, così da renderli poi destinatari, in modo

29 Cfr. C. O’Neil (2016), pp.179-185. Di frequente vengono raccolti indizi che manipolazioni sistematiche siano già in atto. Un esempio riguarda le raccomandazioni che YouTube (Google) sottopone a ogni singolo utente ogni volta che si apre un video e che sono governate da un AAA la cui finalità dovrebbe essere quella di massimizzare il tempo trascorso su YouTube per accrescere il volume di pubblicità. Sulla base dei dati raccolti (attraverso una simulazione) durante la Campagna elettorale americana del 2017 da un ex-programmatore di YouTube, Guillaume Chaslot, il Guardian ha stimato che, considerando i video raccomandati a persone che avevano raggiunto YouTube alla ricerca dei termini “Trump” o “Clinton” e prendendo quelli che mostravano una chiara inclinazione per uno dei due candidati, i video che risultavano a favore di Trump erano ben l’86% del totale. Suggestendo una manipolazione (non provata) dell’algoritmo che presiede alle raccomandazioni (cfr. <https://www.theguardian.com/technology/2018/feb/02/youtube-algorithm-election-clinton-trump-guillaume-chaslot>). Per una “risposta” a queste preoccupazioni, si veda, nel caso di Google, il Documento <https://www.blog.google/around-the-globe/google-europe/fighting-disinformation-across-our-products/>.

30 Si vedano le informazioni fornite dal sito <https://algotransparency.org/?date=16-02-2019&keyword=>. E’ significativo che nel gennaio 2019, a seguito di una vasta campagna di pressione, YouTube abbia annunciato un intervento sull’algoritmo che stabilisce le raccomandazioni per ridurre le raccomandazioni di video con “contenuto ai limiti dell’accettabile (*borderline*) o che potrebbe disinformare gli utenti con esiti dannosi”. Nel febbraio 2019 Google ha pubblicato un Documento su come combattere la disinformazione: <https://www.blog.google/around-the-globe/google-europe/fighting-disinformation-across-our-products/>.

31 Questi quattro scenari sono descritti in <https://render.betaworks.com/synthetic-media-d0adcc53800a>, dove si argomenta che è anche possibile immaginare una combinazione degli scenari descritti e si sollecita a un confronto serrato su di essi.

inconsapevole, di messaggi politici differenziati da parte di un dato candidato. Di nuovo, siamo in presenza di una tattica politica antica quanto il mondo. La novità sta nella scala e nella granularità con cui gli algoritmi consentono di realizzarla, frammentando l'elettorato in minute categorie; e nella sistematica capacità di indirizzare a ognuna di queste categorie, in modo massiccio, un messaggio diverso, ricevuto in modo individuale al di fuori di un contesto collettivo. Questa iper-segmentazione, che esaspera il processo di frammentazione sociale comunque in atto, contribuisce a rendere opaco e indistinto il progetto politico del candidato – che apparirà ad ognuno in modi diversi - e mina il confronto collettivo e quindi il processo di formazione di opinioni condivise.

Il che ci porta al terzo effetto, il più subdolo, che opera quotidianamente e che può spostare il “senso comune” in una direzione che spiazzava il confronto democratico. Questo rischio è racchiuso dentro quello che si presenta come un vantaggio: l'estrazione attraverso gli algoritmi di opinioni e orientamenti politici diffusi dal rumore della rete. Il giudizio positivo su questa funzione deriva dall'assunto che le affermazioni che noi utenti affidiamo alla rete riflettano le nostre opinioni e i nostri orientamenti. Ma è così?

In realtà, il nostro linguaggio quotidiano, il significato delle nostre parole, riflette le nostre distorsioni o pregiudizi istintivi, che sono latenti nella nostra personale cultura³². Ma il nostro comportamento, le nostre opinioni e i nostri orientamenti possono assumere forme che si allontanano da tali pregiudizi istintivi, che tuttavia lasciano traccia nel linguaggio. Questo allontanamento avviene attraverso processi di cui è parte importante il confronto acceso e aperto con opinioni diverse, dove si manifesta la pressione dei sentimenti degli altri sui nostri sentimenti, e viceversa³³. Ma non è questo che gli algoritmi colgono quando elaborano milioni di parole raccolte in rete. Essi colgono piuttosto le distorsioni istintive racchiuse nel linguaggio, che in questa fase storica sono particolarmente forti, viste le disuguaglianze, la rabbia, il risentimento. Ecco allora, come è stato provato, che l'applicazione degli algoritmi al linguaggio ordinario raccolto dalla rete li induce a “incorporare le stesse distorsioni che sono implicite nel linguaggio umano”³⁴.

Ma questo esito ci viene restituito come se si trattasse delle nostre opinioni e dei nostri orientamenti. Ammantato dal crisma di un'elaborazione oggettiva, questo messaggio viene da tutti noi interpretato come l'“opinione prevalente”, il “nuovo senso comune”. L'effetto distorsivo è grave, anche senza immaginare alcuna manipolazione. Insomma, l'uso degli algoritmi per estrarre messaggi dal rumore della rete finisce per sospingere il senso comune collettivo in una direzione che non corrisponde a ciò che emergerebbe da un confronto aperto, informato e ragionevole.

Questi tre effetti sistemici che possono alterare la formazione delle opinioni sono di particolare gravità perché non presumono attività illegali o abusi della rete. Non sono patologie, sono possibili fisiologie dell'uso degli algoritmi nella raccolta e diffusione delle informazioni. Per questa ragione contrastarle è una sfida necessaria e impegnativa.

Una diversa strada è possibile ... ma quale?

Negli anni più recenti è rapidamente cresciuta la consapevolezza dei rischi per la giustizia sociale del “sistema tecnologico delle decisioni automatiche” costituito dagli algoritmi di apprendimento automatico e dalla moltitudine di dati identitari in rete, rischi che, come si è visto, toccano tutte le dimensioni della vita umana. E' cresciuta la consapevolezza che è necessario sottrarsi alla deriva di questo modello centralizzato affidato al controllo di grandi monopoli digitali: una sorta di “capitalismo della sorveglianza” dove la sovranità su molte nostre decisioni di vita è stato trasferito a quei monopoli. E che l'alternativa non possa essere costituita da un altro modello centralizzato, quello della “sorveglianza di stato”, sperimentato dal sistema cinese del *social credit scoring*, con l'attribuzione di punteggi individuali (e penalizzazioni) in relazione al grado di aderenza a norme sociali mostrato dalle singole persone nella vita quotidiana³⁵.

32 Cfr. in particolare Caliskan-Islam, J.J. Bryson, A. Narayanan, *Semantics derived automatically from language corpora necessarily contain human biases*, mimeo, Princeton University-University of Bath, August, 2018.

33 Su questo punto cfr. J.Heidt, *Why Good People Are Divided by Politics and Religion*, Pantheon Books, 2012

34 Cfr. ancora Caliskan-Islam *et al* (2018), p.11.

La reazione si manifesta in una crescente letteratura specialistica sull'argomento, dalla quale abbiamo tratto le precedenti illustrazioni, in studi e analisi e nel moltiplicarsi di corsi sull'etica dell'intelligenza artificiale. L'impegno di ricerca e poi di denuncia ha riguardato lo stesso mondo dell'informatica, con un ruolo particolarmente importante delle donne.

Alla sensibilità della tradizione femminista per il tema del lavoro non pagato o sottopagato, fenomeno così eclatante nell'economia digitale, si accompagna una reazione al processo di maschilizzazione di questo comparto iniziato a inizi anni '80. Fino a quel periodo gli stereotipi di genere avevano agito nel senso opposto, contribuendo ad una forte femminilizzazione delle attività informatiche: la "pignoleria" e "precisione" richiesta negli anni '50 per controllare le righe di codice e immaginare come le macchine le avrebbero eseguite quadrava con l'immagine del "lavoro per le donne"³⁶. Ma allora entra nelle Università e poi nel lavoro una leva in cui i maschi, assai più delle femmine, erano stati indotti, prima nelle famiglie poi dalla cultura, a spendere ore con gli "apparecchi elettronici" e a considerare inadeguato chi non lo faceva³⁷. Fra il 1990 e oggi la percentuale di femmine occupate in ruoli di programmazione informatica scende negli USA dal 35% al 26%; è fra il 17 e il 20% nelle società Facebook, Twitter, Google e Microsoft. Contro uno squilibrio di genere così mercato e le sue conseguenze si osserva oggi mobilitazione di molte donne informatiche, specie nei paesi anglosassoni.

La reazione coinvolge oggi anche figure che hanno svolto un ruolo fondamentale nello sviluppo della rete e degli algoritmi. Ha iniziato a coinvolgere le autorità di regolazione della finanza. A seguito, poi, di vicende eclatanti che hanno toccato la democrazia e la politica, la reazione ha coinvolto le autorità di governo.

L'Unione Europea ha approvato nel 2016 ed è in vigore un Regolamento generale sulla protezione dei dati ed è fortemente impegnata su questo terreno (ci ritorniamo). I principali paesi europei, con l'esclusione dell'Italia, sono in movimento, soprattutto sul terreno delle false informazioni. La Germania, dopo un tentativo non riuscito di autoregolazione, dal gennaio 2018 prevede una sanzione di 20 milioni di euro per le società digitali che non rimuovono entro 24 ore espressioni di odio dal web: il risultato è che Facebook concentra oggi un sesto di tutti i suoi "moderatori" al mondo proprio in Germania. La Francia prevede dal novembre 2018 la possibilità per i giudici di rimuovere articoli che siano considerati "disinformazione". E ha avviato, sempre nel 2018, un'offensiva nei confronti dei giganti del web: prima portando Google e Apple in giudizio con l'accusa di praticare condizioni contrattuali abusive nei confronti di sviluppatori o start-up francesi; poi con l'annuncio che l'Assemblea nazionale e il Ministero della Difesa non utilizzeranno più Google come motore di ricerca. La Gran Bretagna, dopo avere costituito nel 2017 un Centre for Data Ethics and Innovation, ha preso attraverso la Commissione parlamentare per digitale, cultura, media e sport una posizione molto dura su Facebook, sostenendo che materiali interni di quella società dimostrano che la piattaforma ha "intenzionalmente e consapevolmente" violato sia la privacy dei dati, sia le regole della concorrenza.

Sono così cresciute le proposte per correggere l'attuale stato delle cose, per imboccare una strada diversa. Possiamo raggrupparle in due vie di fuga che chiameremo: *Riequilibrare il mercato e organizzare gli utenti per trasformarli in lavoratori remunerati*; *Costruire comunità sicure e chiuse fondate su Blockchain*. E in una strategia: *Sperimentare soluzioni che modifichino il sistema tecnologico esistente*.

Le proposte su cui invitiamo a discutere e lavorare sono quelle racchiuse nella strategia. Ma le prime due vie di fuga ci aiutano a inquadrarla, perché indicano percorsi internamente coerenti che colgono la natura radicale dei problemi da affrontare e in un caso sono di ispirazione per i tentativi in atto.

Via di fuga n.1 Riequilibrare il mercato e organizzare gli utenti per trasformarli in lavoratori remunerati

³⁵ Cfr. Y.J. Chen et al., *Rule of Trust: The Power and Perils of China's Social Credit Megaproject*, 32 *Columbia J. Asian Law* 1, 2018.

³⁶ Cfr. Clive Thompson, *The secret history of women in coding*, febbraio 2019, ripubblicato in *Internazionale*, n. 1297.

³⁷ Idem.

A proporre questa prima via di fuga sono gli esponenti radicali del pensiero liberale, che credono fortemente nella capacità del mercato, purché si tratti di un mercato concorrenziale, dove il lavoro sia organizzato, concorra a impedire posizioni di monopolio ed extra-profitti e sia così anche incentivato a meglio contribuire alla produzione. L'enunciazione di questa opzione ha trovato una forte espressione nel recente volume *Radical Markets* di Eric Posner (Chicago University) e Glen Weyl (Microsoft), nel quale, di fronte all'esplosione delle disuguaglianze, gli autori propongono molteplici riforme radicali.

Il punto di partenza è lo scambio ineguale fra utenti della rete che forniscono dati e imprese che li utilizzano: queste concedono ai primi servizi utili e godibili, ma si impossessano del valore di mercato di quel contributo di lavoro. L'argomento tradizionale che il valore marginale di ognuno di quei dati è infinitamente piccolo (rappresentando un infinitesimo contributo aggiuntivo rispetto a tutti gli altri dati) viene criticato, osservando che ogni contributo di informazione – il “lavoro” di ogni utente – non concorre ad alimentare un dato, unico processo produttivo, l'esecuzione di un dato algoritmo, ma alimenta un intero sistema di intelligenza artificiale dove ogni giorno si eseguono innumerevoli algoritmi e se ne testano altri, sempre più complessi, il cui valore di mercato cresce progressivamente (pp. 226-229). Gli extra-profitti che derivano da questo scambio ineguale – prosegue il ragionamento – non sono dunque giustificati. Mentre, al tempo stesso, viene sottoutilizzato il contributo degli utenti, la loro capacità di fornire in rete un contributo di competenza e idee.

La soluzione proposta sta nel completare il mercato, nel passare, per il sistema tecnologico - intelligenza artificiale + produzione di dati in rete – “dal feudalesimo al capitalismo”, come essi scrivono: remunerare il lavoro di produzione e fornitura dei dati da parte degli utenti/lavoratori della rete. L'esistenza di uno scambio di mercato condurrà alla formazione un prezzo che corrisponderà alla quantità/qualità delle informazioni fornite nella transazione; l'utente/lavoratore potrà così (e sarà incentivato a) far valere le proprie competenze e informazioni. Infatti, nel colloquio virtuale con cui gli autori rappresentano lo scambio – fra l'utente/lavoratore e una voce automatica a nome dell'impresa – l'utente/lavoratore fornisce, a richiesta, informazioni private su sé stesso e sui propri amici.

Perché richiamare questa soluzione? Essa affronta solo uno, pure importante, dei molteplici rischi per la giustizia sociale dell'attuale sistema: lo scambio ineguale e i suoi effetti diretti sulla distribuzione del reddito e della ricchezza e sull'efficienza produttiva. Mentre non tocca tutti gli altri rischi, relativi alle distorsioni e degenerazioni che gli AAA e il nostro mancato controllo sull'uso dei dati identitari possono produrre. Anzi li aggrava, prefigurando un mondo distopico in cui ognuno di noi vende o fabbrica informazioni private su tutti gli altri, consegnandole a piattaforme proprietarie³⁸.

Nonostante ciò, la prima ragione per richiamare questa via di fuga è che essa sottolinea con forza l'iniquità insita nello scambio ineguale, un tema da cui ogni altra soluzione non può prescindere. La seconda ragione è che essa è meno lontana dalla realtà di quanto si pensi: fra il 2016 e il 2018, Facebook ha pagato a giovani fra 13 e 35 anni 20 euro al mese perché “vendessero la propria privacy”, cedendo attraverso un'app (Facebook Research) tutte le informazioni relative all'attività telefonica e su rete³⁹. La terza ragione sta nel modo, nell'unico modo, in cui gli autori si immaginano che a questa soluzione si potrebbe arrivare, stante la forza delle “sirene dei server”: l'emergere di un sindacato mondiale che organizzi noi utenti/lavoratori, mobilitandoci per scioperi/boicottaggi, che interrompano a un tempo la fornitura di dati e il consumo di servizi. Non appare certo uno scenario realistico, ma ci restituisce il senso della potenza delle forze in campo ed è significativo che, all'interno delle avanguardie del pensiero liberale, nella ricerca di un modo per spezzare i monopoli, si torni con coerenza a proporre il ritorno dei sindacati.

Via di fuga n.2 Costruire comunità sicure e chiuse fondate sul Blockchain

³⁸ Su questo punto, cfr. fra gli altri N. Srnicek, *The Social Wealth of Data*, in AA.VV. *Social Wealth Fund*, June 2018.

³⁹ L'app è stata interrotta nel gennaio 2019 quando il team di TechCrunch ha portato alla luce questa vicenda, che tra l'altro violava le regole di Apple: cfr. https://techcrunch.com/2019/01/31/facebook-researchgate/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlMmNvbS8&guce_referrer_cs=OTYK8uYnFfmx2QWTVHIOXg.

Come la prima via di fuga si affida completamente al mercato, esasperandone il ruolo, così questa si affida completamente alla comunità, una comunità che non ha bisogno di relazioni fiduciarie perché è tenuta insieme da un unico destino, attorno alla tecnologia Blockchain.

La prima strategia non mette in discussione la concentrazione dei dati e della loro elaborazione in poche mani private, ma pensa di potere togliere loro i benefici monopolistici attraverso la costruzione di un mercato concorrenziale. Questa strategia, invece, mira a rendere i dati e la loro elaborazione “patrimonio comune”, patrimonio condiviso di una comunità. E’ il punto di partenza e di forza di questa impostazione, che ritroveremo anche nella nostra strategia: l’idea di un patrimonio comune delle informazioni, che sia a disposizione di tutti, con una tecnologia che non dia vantaggi competitivi a nessuno, che garantisca il nostro controllo e la sicurezza dei dati e che favorisca la collaborazione diretta fra utenti, senza passare attraverso alcun “centro” terzo, privato e pubblico che sia. Potremo disegnare e usare algoritmi, prosegue questa tesi, assieme agli altri soggetti che vi concorreranno con i propri dati, e secondo i principi che aderiscono ai valori nostri e degli altri. Si tratta chiaramente di una via di fuga dal presente sistema che cerca il ritorno allo spirito iniziale, anti-statale e anti-capitalista, con cui partì la rete. Ma la comunità con cui si cerca di risolvere il problema è una comunità tutt’altro che universale, è piuttosto una comunità chiusa. E questo è scritto nella tecnologia che adotta.

Il Blockchain o catena di blocchi, è un’architettura di database distribuita: una stessa informazione viene immagazzinata su molteplici nodi, collegati in una rete. L’architettura è basata non sul riferimento a un’informazione originale autoritativa, ma sul consenso tra i nodi della rete. In queste caratteristiche risiedono tre possibilità. Primo: archiviare i propri dati personali, i descrittori della nostra identità, in modo non accessibile se non per nostra scelta. Secondo: realizzare un autogoverno distribuito di questi dati. Terzo: trasferire informazioni e realizzare transazioni fra due o più parti con una garanzia collettiva; un’opzione che consente – è l’impiego iniziale e più noto – di creare moneta all’interno di una comunità di utenti⁴⁰.

Ma nella natura della tecnologia sta anche il duplice limite di questa strategia. In primo luogo, la “comunità” costruita dal Blockchain non si regge sulla “fiducia”; il sistema è affidabile indipendentemente dall’affidabilità dei suoi singoli nodi, perché è il codice condiviso che stabilisce le regole interne. Questo tratto, in linea peraltro con l’ambiente anarco-libertario da cui la tecnologia nasce, mette la comunità al di fuori delle regole “esterne” ed è in questo senso la libera dalla sovranità di poteri privati o statali. Ma al tempo stessa la rende penetrabile da (e quindi attraente per) comportamenti opportunistici. Il secondo limite è legato al primo: la strategia non è adatta a permettere transazioni con l’”esterno”, con soggetti che non fanno parte della Blockchain stessa. Se la comunità decide di realizzare, per l’impiego delle proprie informazioni, un “contratto” con altre parti, essendo coinvolte informazioni replicate in ogni nodo, il contratto deve coinvolgere ogni nodo, senza tuttavia la garanzia che la reazione di ogni nodo sia la stessa. La sola soluzione per aggirare il problema è di affidarsi a un’entità terza di cui fidarsi per incorporare le informazioni esterne nella catena: ma questa strada mina lo stesso obiettivo del sistema decentralizzato⁴¹.

Le obiezioni che vengono spesso mosse a questa soluzione riguardano la sua realizzabilità e sostenibilità. Ci si domanda perché la forza monopolistica dei giganti del web e delle altre imprese che acquisiscono dati gratuitamente e offrono o vendono servizi e che per farlo impiegano algoritmi dovrebbe cedere il passo a queste nuove forme? Perché l’attuale sistema di scambio ineguale non dovrebbe finire per trattenere gli utenti/lavoratori rispetto all’ipotesi di spostarsi sulle tante versioni dell’alternativa? E si sottolinea lo straordinario consumo di energia elettrica che il ricorso al Blockchain (con i suoi lunghi tempi di elaborazione degli scambi) richiede. Ma è piuttosto il duplice problema prima indicato a suggerirne i limiti.

40 Cfr. D. Tapscott, A. Tapscott, *Blockchain Revolution*, Penguin 2018. Gli autori scrivono che con il Blockchain, si può configurare un “autogoverno distribuito del sistema dei dati identificativi delle persone” in cui il modello a cui è affidata la tutela dei dati sia “distribuita fra, e conservata da, le stesse persone tutelate, in modo tale che gli incentivi di tutti sono allineati – un patrimonio comune di dati identificativi – con chiari diritti degli utilizzatori di prendersi cura dei propri dati, accedervi e consentire ad altri di accedervi, trarne ricavi e partecipare alla fissazione delle regole per il mantenimento e l’uso del bene comune” (p. xlvii).

41 Su questo aspetto, cfr. <https://www.the-blockchain.com/2016/04/12/beware-of-the-impossible-smart-contract/>.

Tuttavia, questa via di fuga rappresenta, come la prima, un utile punto concettuale di riferimento, e, ben più della prima, lo sprone per ricercare vie di uscita dallo stato attuale delle cose, e la fonte di sperimentazioni, da seguire e valutare.

Si pensi, nel campo della salute, al caso (riportato da Tapscott&Tapscott) di una rete universitario/ospedaliera che attiva un portale in cui ogni paziente può accedere all'intero materiale della propria storia medica di cui è proprietario e che può a propria discrezione decidere di condividere; prefigurando l'ipotesi di muoversi verso una vera e propria "proprietà dei dati da parte del paziente", alleato con altri – della propria comunità, si intende - e con ricercatori in vere e proprie "cooperative della salute" capaci di esercitare un potere negoziale. O si pensi a una rete come Faircoop (<https://fair.coop/it>) che, utilizzando una versione di blockchain diversa e a minore consumo energetico, promuove e facilita esperimenti di scambio di risorse economiche, di spazi e di tempo di assistenza che tengono conto sia dei bisogni che della condizione economica e sociale dei contraenti. Questo tipo di sperimentazioni ci porta alla strategia da adottare.

Strategia: Sperimentare soluzioni che modifichino il sistema tecnologico esistente

La strategia qui descritta è in realtà un insieme di interventi individuali, collettivi, pubblici e privati che, muovendosi all'interno del "sistema tecnologico" esistente (fatto di algoritmi e dati identitari), convergono nella direzione di piegarlo a esigenze di giustizia sociale. In quanto segue, non si ha alcuna pretesa di "trovare una quadra" fra queste diverse ipotesi, che oggi non è alle viste. Si è piuttosto ritenuto utile raccogliere e proporre all'attenzione della ricerca, dell'azione collettiva e dell'azione pubblica dell'Italia: nonostante singoli e significativi contributi, manca in Italia un confronto pubblico adeguato e una mobilitazione sociale, politica e culturale su questi temi. Il ForumDD ritiene che possa e debba essere compiuto da tutte le parti un salto di qualità.

La strategia si inquadra nell'ambito della regolamentazione europea in tema di protezione dei dati, si articola in un insieme di interventi integrati su piattaforme comuni e sistema tecnologico e dà impulso allo sviluppo delle comunità di innovatori in rete. Questa strategia può trovare un momento di coagulo e forza nelle città, come avanguardia dell'inversione di rotta.

Regolamentazione europea in tema di protezione dei dati

Con il Regolamento generale per la protezione dei dati in vigore dal maggio 2018, l'Unione europea, riconfermando la sua posizione di leader internazionale nel campo della regolazione, ha offerto un quadro di riferimento giuridico che tocca entrambi gli aspetti del "sistema tecnologico" in questione: come disciplinare la raccolta e l'uso dei dati; come regolare i processi decisionali che utilizzano algoritmi di apprendimento automatico. Il contributo di Giorgio Resta al ForumDD (raccolto nei Materiali), analizza ruolo e limiti di questo importante passo.

Il Regolamento definisce i "diritti digitali" in merito al trattamento dei dati personali e alle decisioni che siano assunte attraverso l'utilizzo di algoritmi di apprendimento automatico. Per quanto riguarda l'uso dei dati personali, vengono stabilite condizioni che l'utilizzatore deve rispettare quando tale uso serva a "profilare" una persona (analizzare e prevedere aspetti della sua vita): informazione alla persona; correttezza, incluso un uso non discriminatorio; minimizzazione dei dati usati e del tempo di conservazione dei dati; restrizione delle finalità d'uso a quelle originariamente previste; valutazione dell'impatto dell'uso sulla protezione dei dati stessi.

In tema di decisioni assunte sulla base di algoritmi, due sono le linee di intervento principali: il diritto di ottenere informazioni circa "la logica utilizzata [dall'algoritmo], nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato" (art.15); il "diritto [dell'interessato a] non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla persona" (art.22).

Si tratta di due questioni decisive per l'impatto sulla giustizia sociale delle nuove tecnologie: accedere alla logica delle decisioni, per poterla contestare; preservare una relazione umana nel processo decisionale. Il

Regolamento europeo, individuandole con chiarezza, le mette al centro del pubblico confronto, dell'azione degli Stati e dell'agire in sede giudiziaria e politica. Non può tuttavia dare loro soluzione, come osserva Resta, perché la specificazione di quei diritti è complessa e perché essi entrano in conflitto con altri diritti.

Per quanto riguarda il diritto a informazioni sulla “logica”, esso si scontra (e rischia di soccombere) quando incontra il diritto di tutela della proprietà intellettuale degli algoritmi stessi. E cozza contro il fatto che la “logica” degli algoritmi non è di natura “causale”, derivata da una teoria e come tale contestabile: si potrà al massimo – e già non è poco - comprendere da quali correlazioni è dominata una previsione/decisione, se essa nasconde discriminazioni o errori. Quanto al diritto a relazioni umane, il riferimento agli “effetti giuridici” è limitativo, mentre il divieto può essere aggirato prevedendo l'integrazione delle decisioni dell'algoritmo con una presenza umana, magari non rilevante. Inoltre, lo stesso Regolamento prevede eccezioni al divieto, come nel caso in cui l'affidamento della decisione esclusiva all'algoritmo è stabilita da contratto. Con il “consenso della persona” gli algoritmi possono avvalersi anche di dati particolari altrimenti preclusi e relativi a: salute, orientamento sessuale, opzioni ideologiche e sindacali, appartenenza etnica, etc.

In conclusione, il Regolamento stabilisce una base di principi destinata a fare scuola nel mondo. Non può costituire “la soluzione”, per la natura collettiva e politica dei rischi derivanti dal “sistema tecnologico delle decisioni automatiche” e quindi delle soluzioni, che travalicano, come scrive Resta, la logica dei diritti individuali” del Regolamento. Ma costituisce un quadro di ordine e di riferimento per le altre azioni collettive e pubbliche necessarie.

Interventi integrati su piattaforme comuni e sistema tecnologico

Raggruppiamo qui tutti i diversi interventi, pubblici o collettivi, che, sfruttando il contesto regolatorio (e magari alimentandolo), non assumono come imm modificabili le modalità di raccolta dei dati identitari e gli algoritmi che li utilizzano, ma mirano a valutare per quali obiettivi e con quali valori raccogliere i dati e impiegare gli algoritmi, in quali processi decisionali, con quali diritti riconosciuti a chi i dati fornisce o a chi comunque risente del loro uso. Le strade in questione riguardano “l'intera catena produttiva” necessaria a produrre e aggiornare gli algoritmi, incluso il “lavoro non riconosciuto e sottopagato” necessario a rendere i nostri dati identitari in rete utilizzabili.

All'interno di questa famiglia di interventi, possiamo individuare diverse linee di azione. La prima di esse incalza gli interventi dei giganti del web e delle altre imprese digitali e la ricerca tecnica interna agli algoritmi.

Sotto la pressione di una crescente attenzione agli effetti sociali e alle violazioni della privacy e di altri canoni etici da parte degli algoritmi di apprendimento automatico, i giganti del web – Google, Facebook, IBM, Microsoft, etc - si sono impegnate nella redazione di codici etici o nella costruzione e impiego di meccanismi per la riduzione delle discriminazioni.⁴² Di fronte alla diffusione delle false informazioni e in genere alle preoccupazioni per gli effetti sistemici negativi sulla formazione delle opinioni politiche (di cui si è detto), Google ha di recente sentito il bisogno di descrivere in un documento comprensivo “*How Google fights disinformation*” le sue azioni in questa direzione: accanto a interventi più tradizionali - promozione di “giornalismo di qualità”, metodi per difendersi da attacchi informatici – emerge anche un impegno nella ricerca sui rischi dei *synthetic media*⁴³, ad aprire le proprie ricerche all'esterno, a non dare rilievo a “pagine con contenuti palesemente inaccurati o tesi complottiste” e a contrastare tentativi di manipolazione della gerarchia delle segnalazioni (*ranking system*)⁴⁴. E' un segno positivo, coerente con le generali crescenti preoccupazioni delle classi dirigenti economiche per il risentimento indotto dalle crescenti disuguaglianze. Ma, come sempre, questi passi non possono promettere cambiamenti effettivi se la verifica di “eticità” continua a essere delegata a tecnici senza meccanismi di verifica esterna, garanzie di pubblicità e sanzioni.

42 Cfr. Al Now (2018), pp. 28-32. Nel codice etico di Google fra l'altro si legge l'impegno a “produrre un impatto sociale positivo” e “evitare di creare o rinforzare pregiudizi”.

43 Si legge: “investimenti in ricerca per comprendere come gli algoritmi possano aiutare a filtrare i ‘contenuti sintetici’, lavorando con esperti in tutto il mondo”.

44 Si veda anche la quasi contemporanea decisione di YouTube descritta prima descritta.

L'allerta delle imprese e la pressione del confronto, anche all'interno del mondo della ricerca, hanno comunque promosso lo sviluppo di metodologie volte a tenere conto di obiettivi di giustizia sociale negli algoritmi. Le sperimentazioni e il dibattito sulla loro efficacia è in corso. Se ne segnalano due conclusioni particolarmente importanti. La strada di escludere dagli algoritmi le variabili che direttamente o indirettamente appaiono correlate con caratteristiche delle persone tradizionalmente sfavorite (per genere, gruppo etnico, ceto sociale) – o *anti-classification strategies* – danno risultati insoddisfacenti: su questa strada – è la conclusione che va emergendo - è meglio riconoscere espressamente il valore da tutelare, prevedendo per quelle caratteristiche soglie diverse.⁴⁵ Se poi, come si è visto⁴⁶, l'applicazione degli algoritmi al linguaggio ordinario raccolto dalla rete li porta a incorporare le stesse distorsioni che sono implicite nel linguaggio umano e che non necessariamente corrispondono ai nostri convincimenti, allora la strada non può essere quella di correggere gli algoritmi. Si tratta piuttosto di utilizzarne i risultati non come “decisioni” da adottare, ma come input di un processo decisionale da affidare ad altre valutazioni e strumenti che prevedano il confronto aperto di opinioni diverse.

E' dunque di primaria importanza esercitare una pressione sociale, politica e istituzionale sui giganti della rete e sulle altre imprese rilevanti nell'impiego degli algoritmi affinché *aprano le loro piattaforme alla ricerca esterna*, rendano disponibili e verificabili esiti e materiali della ricerca interna, modifichino gli algoritmi e rendano monitorabili tale modifiche, accettino in genere una verifica della loro aderenza agli impegni etici assunti, e in prospettiva affinché rendano gli algoritmi *open source*. Ma la modifica dell'impatto sociale del sistema tecnologico può avvenire solo a seguito di interventi che tocchino il sistema intero, sia sul piano giuridico che sul piano delle azioni collettive. Veniamo dunque alle altre linee di azione della strategia:

1. Chiedere e promuovere una *composizione diversificata dei gruppi di ricerca* che elaborano gli algoritmi. Al fine di pesare sulla scelta di valori e obiettivi e dei dati, nei gruppi di ricerca deve essere assicurata un'“adeguata rappresentazione di punti di vista diversi”⁴⁷ per disciplina, esperienza, sensibilità a temi di giustizia sociale, con ottica nazionale e internazionale. Di particolare importanza, per i motivi detti, è un radicale riequilibrio di genere. E' il primo passo che rende evidente ed effettivo che gli algoritmi non sono “tecniche” oggettive.
2. Agire affinché la costruzione di ogni progetto di ricerca e di ogni algoritmo sia *aperta al confronto pubblico e informato* con le persone che producono i dati impiegati o che risentono delle decisioni a cui gli algoritmi conducono (lavoratori, consumatori); e affinché tali decisioni siano anche esse oggetto di confronto. E' questo l'obiettivo di Istituzioni come la citata AI Now negli USA o come Decode in Europa - di cui fa parte Il Politecnico di Torino – che mira allo “sviluppo di strumenti che diano alle persone controllo sui dati che forniscono per uso privato e collettivo” e come altre ancora⁴⁸.
3. Costruire e utilizzare *piattaforme digitali che ridiano “sovranità tecnologica” agli utenti/lavoratori* della rete e in genere ai cittadini toccati dalle decisioni assunte sulla base dell'analisi dei dati. Come costruire piattaforme digitali comuni è al centro di un confronto con diverse impostazioni⁴⁹ che hanno in comune, rispetto al modello Blockchain, la natura delle comunità che si avvalgono della piattaforma: comunità aperte e basate su relazioni di fiducia.
4. *Moltiplicare la quantità dei dataset aperti*, requisito indispensabile di piattaforme digitali comuni. Gli Stati nazionali e l'Unione Europea potrebbero dare un contributo ben superiore in questa direzione, tenendo conto che molti di questi dataset sono stati pagati dai contribuenti: dati catastali, dati sui trasporti pubblici, identificativi unici delle imprese, dataset legati a numeri civici.

45 Cfr. J. Kleiberg, J. Ludwig, S. Mullainathan, A. Rombachan, *Algorithm Fairness*, AEA Proceedings, 2018, 108.

46 Cfr. Caliskan-Islam *et al* (2018).

47 Cfr. in particolare R.Abebe, *Why AI needs to reflect society*, Forbes Insight, nov. 2018. I principi da lei sostenuti sono attuati dal Gruppo di ricerca Mechanism Design for Social Good da lei co-fondato.

48 Cfr. la descrizione di esempi in F.Bria, E.Morozov (2018), pp. 31-32.

49 Altre proposte di costituzione di Piattaforme collettive dei dati identitari sembrano viceversa fare affidamento su un “centro pubblico”: è il caso della proposta di un “*National Data Fund*” avanzata da N. Srnicek (2018). Anche in questa proposta sono le persone a decidere se condividere i propri dati, secondo “vincoli granulari” sul loro uso, ma tali dati sono anonimizzati. Proposte o tentativi di costruire piattaforme digitali che ridiano “sovranità tecnologica” agli utenti/lavoratori emergono in forme diverse nel confronto in corso: cfr. ad esempio E. Morozov, F. Bria, *Rethinking the Smart City*, Rosa Luxemburg Shiftung, January 2018.

5. Realizzare azioni collettive e mettere sotto pressione le pubbliche autorità affinché venga data piena attuazione alla *Regolamentazione* europea, sfruttandone tutti gli spazi, e promuovere, ove necessario, nuova regolamentazione.
6. *Promuovere e rimuovere gli ostacoli che prevengono le sperimentazioni realizzate dalle “comunità di innovatori”* che già oggi mettono in atto forme alternative di gestione e governo dei dati identitari.
7. *Realizzare campagne di sensibilizzazione dei cittadini* e delle diverse categorie toccate dalla monopolizzazione delle conoscenze (lavoratori, consumatori, richiedenti credito, assicurati, etc.) circa l'entità, le cause e le conseguenze di tale monopolizzazione, e circa i modi per contrastarla. *Formare* a questi stessi contenuti e soluzioni alternative i giovani studenti sin dai primi anni di scuola. Sono questi requisiti indispensabili affinché si crei una massa critica adeguata ad attivare gli strumenti precedenti.

L'insieme di queste iniziative configura i tratti di un modello di “sovranità collettiva” sui dati personali e gli algoritmi. Un modello che aspira a essere alternativo sia al modello-USA della “sovranità dei monopoli” sia al modello-Cina della “sovranità dello Stato”. Non si tratta, a differenza di questi due, di un modello compiuto. Ma esso trova la sua cornice di riferimento nel punto di forza dell'Europa: essere andata più avanti degli altri nel costruire una cornice di regole. A esse si è aggiunta nel dicembre 2018 la strategia lanciata dalla Commissione Europea nel campo dell'intelligenza artificiale che comprende, tra le altre iniziative: *i)* il *Progetto di orientamenti etici per l'intelligenza artificiale*, che ha proposto una serie di linee guida coerenti con il precedente gruppo di azioni raccolte attorno al principio di un'intelligenza artificiale consapevole dei rischi e valutata in base agli effetti su diritti fondamentali, valori sociali e principi etici, e *ii)* gli orientamenti sull'interpretazione della direttiva sulla responsabilità per danno da prodotti difettosi alla luce dell'evoluzione tecnologica. Ora si tratta di sapere dar vita a queste regole e a questi indirizzi.

Le diverse tipologie di azione volte a sperimentare una sovranità collettiva possono trovare alimento finanziario, promozione e un punto di raccordo nell'iniziativa lanciata dalla Commissione europea stessa nell'autunno 2016, denominata Next Generation Internet. Il suo obiettivo è di influenzare il futuro di Internet nella direzione di una piattaforma interoperabile che incorpori i seguenti principi: apertura, protezione della privacy e dei dati personali, cooperazione, decentralizzazione, controllo in mano agli utilizzatori, la persona al centro. A questi principi non corrispondono, nell'iniziativa, una diagnosi o una chiara strategia. Ma esiste la volontà di ricercarle attraverso consultazioni pubbliche, progetti di ricerca (in particolare nel progetto NGI Forward) e comunità on line di discussione.

Sviluppo di comunità di innovatori in rete

In questa ricerca di soluzioni un ruolo significativo può essere svolto dalle “comunità di innovatori in rete”⁵⁰. Con questa espressione intendiamo forme neo-mutualistiche di impresa che utilizzano tecnologie di rete per produrre piattaforme e serbatoi di conoscenza comune. Questa conoscenza consente a soggetti esterni di “entrare in azienda” e viene utilizzata da essi come un bene pubblico per raggiungere propri obiettivi e per soddisfare bisogni e aspirazioni con soluzioni innovative. In tal modo queste comunità sfidano la logica della proprietà privata delle idee e svolgono di fatto una funzione concorrenziale all'utilizzo monopolistico delle nuove tecnologie; nelle biforcazioni del cambiamento tecnologico possono aiutare ad imboccare strade che accrescono la giustizia sociale, anziché ridurla.

Si tratta di miriadi di innovatori/inventori – hackers o innovatori militanti, convinti che le idee non possano essere oggetto di proprietà privata - che spesso lavorano in piccole comunità (*crews*), e che sfruttano le possibilità di collaborazione su rete in comunità virtuali e la disponibilità e possibilità di elaborazione di grandi masse di dati. Si pensi, nel campo della salute, a chi sviluppa e offre in open-source laboratori per produrre farmaci a prezzi stracciati, o a esperienze di “scienza partecipativa”, dove assieme agli scienziati sono coinvolti utilizzatori o cittadini con forme sostanziali di apporto. I principali “digital common” sono stati prodotti da comunità di innovatori: Wikipedia/Wikidata, il codice dei grandi progetti open source come GNU/Linux, Apache, MySQL o le piattaforme cognitive come StackOverflow. Particolarmente significativo è il caso di OpenStreetMap, ha fornito una mappa del pianeta aggiornabile in tempo reale alle organizzazioni

⁵⁰ F.Bria, E. Morozov (2018) le definiscono “grassroot communities of innovators”.

che si occupano di disaster response, e che hanno bisogno di ridisegnare rapidamente le mappe in seguito alle devastazioni operate da terremoti, inondazioni e tsunami.

Lo sviluppo di queste comunità è ostacolato dal potere dei monopoli. Ma anche da fattori sui quali si può agire. Le comunità di innovatori avrebbero maggiori opportunità di sviluppo se l'Unione Europea, in coerenza con i propositi espressi nella comunicazione “Verso uno spazio comune europeo dei dati”⁵¹ dell'aprile 2018, e i suoi Stati membri investissero in modo massiccio nel rilascio di banche dati aperte; offrirono a queste comunità luoghi fisici in cui operare; promuovessero le loro possibilità di dialogo con le PMI.

Le città come avanguardia per la shared economy

La terza strategia ora descritta può trovare un punto di forza e di attuazione nel governo delle città. Esiste qui l'opportunità, come osservano F. Bria e E. Morozov (2018), di realizzare un'inversione di rotta nella logica delle “*smart cities*”, concepite a lungo come strategie sostanzialmente affidate al governo delle imprese⁵², per ridare invece potere ai cittadini, sia attraverso l'amministrazione pubblica, sia direttamente: *shared cities*. Di questa riappropriazione dei poteri decisionali è parte rilevante la riappropriazione da parte dei cittadini/utenti dei loro dati identitari e dell'uso che ne viene fatto attraverso gli algoritmi di apprendimento automatico.

Le città costituiscono un mercato decisivo per lo sviluppo dei nuovi servizi fondati sull'intelligenza artificiale. Che tocca molteplici campi: mobilità e traffico, energia e illuminazione, incendi, evacuazioni, e prevenzione delle esondazioni, casa e manutenzione degli edifici, rifiuti, spazi pubblici, sicurezza e sorveglianza, etc.. Sono campi in cui, come osservano ancora Bria e Morozov, sono attive in modo crescente grandi imprese tecnologiche, come Siemens, IBM, Cisco, Phillips con un potenziale importante di innovazione dei prodotti e di miglioramento della qualità di vita. Ma, ancora una volta, si tratta di orientare questo cambiamento tecnologico in modo che non benefici solo né principalmente i ceti forti, ma che benefici anche, anzi principalmente, i ceti deboli. Sta qui l'importanza di fare pesare le aspirazioni e la voce di questi ceti sulle decisioni. A cominciare dall'uso che viene fatto dei dati che essi stessi forniscono. Nell'interesse di chi si ridisegna il sistema di mobilità? O si governa lo sviluppo degli affittacamere o di nuove forme di trasporto? O si modificano i sistemi di sicurezza? Tutte questioni dove è centrale e cresce il ricorso agli algoritmi.

Bria e Morozov riconoscono che sul cambiamento pesano scelte nazionali e internazionali (relative ad altre tematiche toccate dalle proposte del Rapporto del ForumDD). E arrivano con prudenza a scrivere che questa linea di azione nelle città “è un modo per prendere tempo”, mentre si cercano soluzioni più generali. Ma è un tempo che può essere fruttuoso. La dimensione “città”, così come le dimensioni sub-cittadine, offrono un terreno dove si possono sperimentare nuove soluzioni attraverso una ricostruzione di percorsi di democrazia deliberativa che coinvolgano, oltre alle imprese, l'amministrazione, i singoli cittadini e le organizzazioni di cittadinanza attiva in cui una parte dei cittadini si ritrova. In questi percorsi possono svolgere un ruolo importante proprio le “comunità di innovatori in rete” che praticano forme alternative al “sistema tecnologico” dominante. Non a caso, Bria e Morozov indicano come obiettivo primario la sperimentazione di forme che modifichino “il regime di proprietà dei dati”.

In questa chiave si possono interpretare le innovazioni principali introdotte, con il contributo della stessa Bria come *Chief Technology Officer*, nella città di Barcellona. E la rete di città - di cui fa parte Milano, all'avanguardia in Italia su questo terreno - che è nata attorno alla condivisione della stessa strategia e alla *Declaration of Sharing Cities*.

⁵¹ Fra le misure previste nella comunicazione, la Commissione ha evidenziato, fra le misure previste, l'intenzione di continuare a sostenere l'impiego di una vera e propria infrastruttura per dati aperti e la creazione di un centro Europeo di supporto per la condivisione dei dati, da realizzare nel 2019

⁵² Cfr. anche F.Barca, *Alternative Approaches to Development Policy: Intersections and Divergences*, in OECD Regional Outlook, 2011, cap. 11.