

Sistemi di decisione algoritmica e disuguaglianze sociali: le evidenze della ricerca, il ruolo della politica

Biagio Aragona

RPS

Aziende e pubbliche amministrazioni usano con sempre più frequenza algoritmi per prendere decisioni sulle vite di clienti, cittadini, utenti e pazienti. Numerosi studi hanno dimostrato che i sistemi di decisione automatizzata possono generare, e soprattutto amplificare, le disuguaglianze sociali. Con questo articolo si intende promuovere la consapevolezza dei rischi associati ad un uso acritico degli algoritmi, ma allo stesso tempo, incoraggiare all'impiego informato dei sistemi di decisione automatizzata. L'obiettivo principale è di portare evidenze empiriche riguardo al funzionamento degli algoritmi e delle loro conseguenze sulla società. Successivamente, si individuano delle tecniche di ricerca sociale

che possono essere impiegate per effettuare audit di algoritmi. Infine, si focalizzano delle soluzioni politiche per mitigare gli effetti degli algoritmi sulle disuguaglianze: la costituzione di organismi di sorveglianza in raccordo con altre autorità garanti (es. comunicazione e privacy), l'alfabetizzazione della popolazione alla privacy dei dati e al coding, la formazione di figure interdisciplinari altamente specializzate capaci di valutare la qualità degli algoritmi e i loro effetti. La ricerca sociale e l'istituzione di organismi di controllo sono gli strumenti principali per governare l'autorità algoritmica e per evitare che si producano nuove, e sempre più complesse, forme di disuguaglianza sociale.

1. Introduzione

Aziende e pubbliche amministrazioni usano con sempre più frequenza algoritmi per prendere decisioni sulle vite di clienti, utenti, cittadini e pazienti che incidono su, ad esempio, la scelta del prodotto da proporre, la concessione di un prestito, l'assegnazione di un posto di lavoro, l'erogazione di una certa quantità di farmaco. Numerosi studi hanno dimostrato che i sistemi di decisione automatizzata possono talvolta generare, e molto spesso amplificare, disuguaglianze sociali (Eubanks, 2017; Noble, 2018; O'Neil, 2016).

In questo programma di ricerca, che intreccia disuguaglianze sociali, mutamento tecnologico e ricerca sociale, si colloca questo articolo; con

l'intenzione di promuovere la consapevolezza dei rischi associati ad un uso acritico degli algoritmi, ma allo stesso tempo, incoraggiare l'impiego consapevole dei sistemi di decisione automatizzata. L'obiettivo è di portare evidenze empiriche riguardo al funzionamento degli algoritmi, delle loro conseguenze sulla società in generale e su alcune fasce specifiche della popolazione. Innanzitutto, si presentano alcuni esempi per individuare gli impatti possibili delle decisioni automatizzate e anticiparne i rischi. Successivamente, si individuano delle tecniche di ricerca sociale che possono essere impiegate per l'analisi dei complessi sistemi socio-tecnici di decisione automatizzata. Infine, si suggeriscono delle soluzioni politiche affinché la diffusione degli algoritmi nelle tante sfere della società non sia veicolo di ulteriori disuguaglianze.

I sistemi di decisione automatizzata, o processi di decisione controllati algoritmicamente, sono dei sistemi procedurali in cui le decisioni vengono inizialmente, in parte o del tutto, delegate a un ente o a un'azienda, che a sua volta utilizza algoritmi per compiere un'azione. Questa delega – non della decisione stessa, ma dell'esecuzione – prevede un modello decisionale, un algoritmo che traduce questo modello in codice calcolabile e dei dati che questo codice utilizza come input. Gli algoritmi, pertanto, nascerebbero dopo aver formalizzato un problema e tradotto gli obiettivi in termini computazionali.

Se a lungo sono stati considerati meri strumenti tecnici per svolgere alcune istruzioni (Sipser, 2006), oramai sono diversi gli studiosi che si discostano da questa visione tecnicista, a favore di un'idea di algoritmo come costruito socio-tecnico, che viene cioè creato da complesse interazioni tra attori umani e «attanti» tecnologici (Latour, 1987). Gli algoritmi vanno ad intrecciarsi con pratiche sociali e materiali che hanno una propria natura culturale, storica e istituzionale (Takhteyev, 2012; Dourish, 2016), influenzando ed essendo influenzati dal contesto in cui vengono impiegati. In questa prospettiva socio-tecnica, l'azione decisionale degli algoritmi viene interpretata come parte di assemblaggi molto più ampi, in cui sistemi di pensiero, conoscenze pregresse, economia, politica, infrastrutture, istituzioni e relazioni interpersonali si intrecciano profondamente su più livelli (Gillespie, 2014; Kitchin e Dodge, 2011). Proprio su queste premesse si sono realizzati numerosi studi che vanno sotto l'etichetta di *critical algorithm studies*, che si concentrano su come le decisioni automatizzate possano amplificare *bias* e disuguaglianze. Questi studi contestano con forza che gli algoritmi, e i dati su cui vengono eseguiti, siano strumenti obiettivi, imparziali e affidabili *by design*, ribadendo insistentemente che:

- gli algoritmi non prendono decisioni, ma eseguono decisioni;
- gli algoritmi, e i dati che elaborano, non sono neutri;
- gli algoritmi, e i dati che elaborano, non sono oggettivi.

2. Algoritmi di disuguaglianza

Il legame tra tecnologia e disuguaglianze è stato a lungo affrontato da diverse prospettive. Una prima serie di argomentazioni riprende la questione dell'uso diversificato delle tecnologie da parte di gruppi sociali con diversa istruzione, età, classe, genere e etnia. In linea con la letteratura sul tema del *digital divide*, si sostiene che i gruppi svantaggiati abbiano minori probabilità di accedere alla rete a banda larga e alle tecnologie digitali, pertanto sviluppino meno competenze e siano allo stesso tempo meno consapevoli dei rischi della tecnologia.

La letteratura sul *digital divide* è stata criticata per la tendenza a separare il sociale dal tecnologico, sottovalutando che è all'incrocio tra svantaggio sociale e scarso accesso alle tecnologie che si possono annidare le disuguaglianze sociali digitali (Halford e Savage, 2010). Il termine disuguaglianze sociali digitali ha proprio il merito di affermare la natura bidirezionale della relazione tra svantaggio sociale e scarsità di accesso alle tecnologie, che invece è sempre stata considerata unidirezionale, perché è lo svantaggio sociale che porta ad un accesso limitato. Esso ci ricorda che fattori sociali strutturali e uso delle tecnologie digitali si configurano a vicenda, si co-costituiscono (Lupton, 2015). Quando si incrociano svantaggi sociali, insufficiente alfabetizzazione digitale, scarse conoscenze dei codici e un accesso limitato alla tecnologia e ad internet, gli individui possono essere maggiormente propensi, da un lato, a sottovalutare i rischi dei sistemi di decisione automatizzata e, dall'altro, ad avere poca fiducia nelle tecnologie algoritmiche.

Ma in che modo gli algoritmi possono generare o amplificare le disuguaglianze sociali?

Airoidi (2020), a partire da una rassegna della letteratura statunitense sul tema, evidenzia una prima questione nel modo in cui gli algoritmi implementati nei motori di ricerca e nelle diverse piattaforme on-line tendono a proporre i contenuti più diffusi (Broussard, 2018). Gli algoritmi sarebbero dotati dell'«autorità» (Rogers, 2013) di influenzare quali fonti siano da considerare più importanti e rilevanti. Questa è, ad esempio, la logica implementata nel motore di ricerca di *Google*, in piattaforme come *YouTube*, e in assistenti virtuali come *Alexa*.

Tuttavia, le conseguenze sociali più pesanti della diffusione della cultura algoritmica (Hallinan e Striplas, 2016) si verificano quando gli algoritmi vengono utilizzati con fini normativi per la gestione dei servizi pubblici. Difatti, i sistemi di decisione automatizzata sono impiegati sempre più spesso nella pubblica amministrazione, a volte con ottimi risultati, a volte con risultati non proprio incoraggianti, ponendo inquietanti interrogativi sul piano dell'equità. Eubanks (2017), ad esempio, dimostra che negli Stati Uniti d'America molte richieste per aiuti sanitari, alimentari o economici sono state negate a causa di sistemi informatici fallaci, indici di disagio inaffidabili e indicatori non validi. L'adozione di sistemi *data intensive* di selezione della platea dei beneficiari delle misure di welfare, che è stata sostenuta da logiche neoliberiste come antidoto all'inefficienza e agli sprechi, ha avuto in alcuni casi conseguenze pesanti sulle vite dei cittadini più poveri e socialmente esclusi, soprattutto afroamericani. In uno dei casi descritti dalla Eubanks nel suo volume, riguardo all'assegnazione dei posti letto ai senza fissa dimora della città di Los Angeles, l'algoritmo andava a contare le notti passate in prigione come *housing*, abbassando l'indice di vulnerabilità a chi veniva arrestato, riducendo così la possibilità di questa persona di accedere in futuro ai pochi posti letto disponibili (*ivi*, p. 126).

In Italia, i sistemi di decisione automatizzata sono diventati tristemente famosi quando il Miur, nell'ambito della riforma educativa della «Buona scuola», decise di assegnare le cattedre per l'anno scolastico 2016/2017 impiegando un algoritmo sviluppato da HP Italia e Finmeccanica. L'algoritmo avrebbe dovuto assegnare la migliore sede possibile ad ogni candidato utilizzando tre criteri principali: l'esperienza lavorativa e i ruoli, le quindici destinazioni preferite dal candidato e i reali posti disponibili. Nell'attribuire le cattedre ai docenti, l'algoritmo avrebbe dovuto anche riconoscere una priorità a coloro che avevano il punteggio più alto in graduatoria. Tuttavia, in circa 10.000 casi (Zunino, 2019) l'algoritmo prese in considerazione unicamente le preferenze espresse dai candidati, senza fare un confronto tra i punteggi e le destinazioni. Non basandosi su punteggi e graduatorie, inviò insegnanti pugliesi e docenti di Catanzaro in provincia di Milano, quando avrebbero dovuto essere destinati alle loro regioni; altri li sottrasse alla loro naturale sede di Padova e, in maniera del tutto incomprensibile, spedì a Prato due professori calabresi con i figli autistici. Quell'algoritmo ha dato il via a migliaia di cause e ricorsi, successivamente vinti in quanto una perizia tecnica in sede giudiziale – richiesta dai sindacati ed eseguita dall'Università di Tor Vergata – definì l'algoritmo «confuso, lacunoso, ampol-

loso, ridondante, elaborato in due linguaggi di programmazione differenti, di cui uno risalente alla preistoria dell'informatica, costruito su dati di input organizzati e gestiti in maniera sbagliata» (Salvucci e al., 2017, p. 12). Il caso della Buona scuola ha portato all'attenzione dell'opinione pubblica che la «governance algoritmica» (Musiani, 2013), basata su processi automatizzati, è un tipo di potere non sempre trasparente. Le sentenze del Tar 8472, 8473 e 8474 del 13 dicembre 2019 hanno confermato che l'algoritmo effettuò decisioni non giustificabili in base ai criteri fissati nell'ordinanza ministeriale n. 241/2016 che attuava il piano di reclutamento straordinario. Con quelle sentenze si è sancito in modo definitivo che le pubbliche amministrazioni italiane possono sì avvalersi di algoritmi per gestire decisioni, ma che ciò è lecito purché l'algoritmo sia reso pienamente conoscibile, cioè trasparente. In caso di errori o disservizi, responsabilità e imputabilità spettano al titolare del potere, che dovrà verificare logica e legittimità della scelta procedurale.

Gli esempi di pubbliche amministrazioni che impiegano sistemi di decisione automatizzata che hanno generato disuguaglianze potrebbe continuare, ma a questo punto la domanda dovrebbe essere chiara: come possiamo noi ricercatori sociali intervenire su questo legame tra decisioni automatizzate e disuguaglianze? Innanzitutto, il punto di partenza è che si deve mettere da parte ogni pregiudizio di tipo ideologico e produrre evidenze empiriche riguardo al funzionamento degli algoritmi e delle loro conseguenze sulla società. Per cui, un algoritmo non può essere considerato buono o cattivo di per sé, ma bisogna entrare nel merito di ogni specifico sistema di decisione automatizzata, del suo assemblaggio e del suo funzionamento. È solo attraverso la ricerca empirica che possono essere compresi gli impatti, anticipati i rischi e le problematiche di natura etica, suggeriti gli interventi per minimizzarli e assicurarsi che, laddove esistenti, buone pratiche trasparenti vengano sviluppate e implementate. Decostruire la natura degli algoritmi, e il modo in cui essi eseguono il potere che gli è stato delegato, è l'unica strada percorribile per analizzare il nesso tra algoritmi e disuguaglianze (Kitchin, 2017).

3. Fare ricerca in ambienti densamente tecnologici: l'audit di algoritmi

Analizzare la multidimensionalità degli assemblaggi di algoritmi significa esaminare diversi domini: gli oggetti materiali (infrastrutture, piattaforme

e, più in generale, sistemi tecnologici), i linguaggi (codici, regolamenti, istruzioni, ecc.) e gli attori sociali che sono direttamente o indirettamente coinvolti nell'assemblaggio (tecnici, utenti, committenti, ecc.). Come direbbero gli studiosi di scienza e tecnologia bisogna aprire la «scatola nera» (Pasquale, 2015) e ricostruire il complesso di dinamiche e di interazioni tra i diversi attori (umani e non umani) che co-partecipano alla costruzione dell'algoritmo.

Una strategia per fare ricerca in questi ambienti densamente tecnologici è ciò che potremmo chiamare un audit di algoritmi (Sandvig e al., 2014; Aragona, 2020). L'audit degli algoritmi intreccia la metodologia della ricerca sociale, l'informatica e l'etica dell'automazione. È un campo interdisciplinare che non ha ancora raggiunto l'autonomia né come oggetto di ricerca, né come tecnica di ricerca. L'audit di algoritmi si concentra su almeno quattro aspetti:

- la qualità dei dati (la loro validità e fedeltà, la loro copertura rispetto alla popolazione di riferimento) e l'esplicitazione delle operazioni di *data curation*;
- la definizione operativa di «successo» dell'algoritmo;
- le scelte tecniche: tipo di algoritmo, tipo di modello, metodi di calibrazione e ottimizzazione;
- la stima delle conseguenze intenzionali e non intenzionali delle decisioni eseguite attraverso gli algoritmi.

Le tecniche per realizzare un audit di algoritmi si distinguono in due grandi famiglie. Le prime si impiegano nel caso i codici, il modello, la documentazione e i dati di partenza siano disponibili (*white box*), le seconde quando non lo siano (*black box*). Per effettuare un audit di algoritmi si può lavorare direttamente sui codici, oppure si possono utilizzare le tecniche tipiche della ricerca sociale.

Tra queste tecniche ci sono sicuramente gli esperimenti, che servono a testare l'algoritmo con un insieme di dati equivalente, oppure in contesti diversi, o a cambiare una piccola parte del codice per vedere che conseguenze procura sui risultati. Un recente esperimento condotto da Kayser-Bril (2020) ha dimostrato che l'algoritmo utilizzato da *Google Vision Cloud* – un servizio di classificazione automatica delle immagini – produceva risultati diversi in base al colore della pelle del soggetto presente nell'immagine. L'immagine da classificare nell'esperimento era una mano che impugnava un termometro palmare, un oggetto che prima della recente epidemia di nuovo coronavirus era conosciuto solo agli specialisti, ma che oggi è diventato un oggetto comune, che gli individui hanno visto in tv, o utilizzare in aeroporti, stazioni, aziende e

ovunque sia diventato obbligatorio rilevare la temperatura. Il sistema di classificazione automatica delle immagini di *Google* era stato istruito su set di dati che probabilmente non contenevano immagini di termometri palmari, pertanto non riuscivano a classificare l'oggetto correttamente. *Google Vision Cloud* andava però a classificare l'oggetto di un individuo nero che impugnava il termometro palmare come «pistola», mentre nella stessa immagine riferita ad un individuo bianco l'oggetto veniva classificato come «strumento elettronico». La causa di questo errore è semplice, se si conosce il modo in cui gli algoritmi «apprendono» dai dati come eseguire delle decisioni. La classificazione viene fatta sulla base di modelli che aiutano la macchina ad interpretare l'immagine, per cui ad esempio, delle mucche che sono fotografate su una montagna innevata, possono essere più facilmente classificate come lupi, mentre, se sono in città, come cani. Il contesto, in pratica, può influenzare la classificazione. Kayser-Bril ha concluso che nelle immagini usate da *Google* per addestrare l'algoritmo, probabilmente gli individui neri erano più frequenti in scene di violenza, pertanto l'immagine di un individuo nero veniva con maggiore probabilità classificata con un termine che appartiene al lessico della violenza. Le conseguenze reali di un *bias* algoritmico così evidente possono essere gravi. Negli Stati Uniti d'America strumenti di riconoscimento automatico delle armi vengono utilizzati nelle scuole, negli stadi e nei supermarket. Così pure in Europa vengono impiegati da alcune forze di polizia nei luoghi affollati. Siccome questi sistemi sono molto simili a *Google Vision Cloud*, potrebbero soffrire lo stesso problema, attribuendo agli individui neri una maggiore probabilità di essere considerati pericolosi. L'esperimento di Kayser-Bril ha indotto *Google* il 6 aprile 2020 a modificare il suo algoritmo, e da allora i termometri palmari vengono classificati correttamente a prescindere dal colore della pelle del soggetto che li impugna.

Anche le tecniche qualitative possono essere molto utili nell'audit degli algoritmi. L'etnografia, ad esempio, si concentra su come gli algoritmi sono generati e come le decisioni pre-analitiche sulla preparazione dei dati e del modello possono avere effetti sui risultati. Mira alla comprensione dei valori simbolici, culturali e normativi che sono racchiusi negli algoritmi, e che promuovono certe immagini della realtà sociale. Ad esempio, Aragona e Felaco (2019) hanno analizzato come veniva sviluppata dal *Web Science Institute* dell'università di Southampton una piattaforma di *datawarehouse* da utilizzare nel sistema sanitario britannico per il monitoraggio di disturbi di natura psichica. Intervistando una serie di stakeholder che partecipavano al progetto, gli autori hanno individuato

RPS

Biagio Aragona

il sistema di attori sia umani (code team e sviluppatori, esperti di dominio, utenti e *stakeholders*), sia non umani (piattaforme, hardware, software, ecc.) che agivano e co-partecipavano alla formazione degli algoritmi che dovevano segnalare i casi più a rischio. L'indagine ha confermato che la costruzione degli algoritmi per il calcolo degli indicatori del cruscotto veniva influenzata dall'azione, dalle esigenze e dagli interessi degli attori che appartenevano in maniera diretta o indiretta all'assemblaggio. Ad esempio, alcuni intervistati hanno affermato di dover calibrare e, in alcuni casi, intervenire sulle istruzioni dell'algoritmo affinché si avvicinasse alle richieste degli psicologi in termini di usabilità e di risultati attesi, oppure che le esigenze di produrre delle esportazioni in formati dati più diffusi tra alcune categorie di studiosi (ad esempio, il software Spss) avevano comportato un intervento sui database e sugli algoritmi di interrogazione e esportazione. Anche se questo studio non ha evidenziato disuguaglianze, è servito a riaffermare la natura co-costitutiva degli algoritmi, che è intrisa di valori, conoscenze di sfondo e culture: tutti aspetti che vengono generalmente espunti dalle retoriche che promuovono l'adozione diffusa dei sistemi di decisione automatizzata, le quali, invece, insistono sull'efficienza, la neutralità e l'obiettività degli algoritmi.

Il novero delle tecniche che possono essere utili per fare audit di algoritmi potrebbe continuare. Ad esempio, con le indagini di tipo survey sugli stakeholder e sulla popolazione oggetto di algoritmo, che hanno lo scopo di rilevare comportamenti e opinioni delle persone interessate da specifiche decisioni automatizzate; come pure con la ricerca partecipativa, in cui individui volontari approfondiscono da utenti il funzionamento di un algoritmo. Quest'ultima tecnica utilizza strumenti di *crowdsourcing*, come ad esempio *Amazon Mechanical Turk*, per individuare soggetti che effettuano ricerche sul web da punti diversi del mondo, e raccolgono i risultati da inviare poi ai ricercatori (Rogers, 2013).

In pratica, sono oramai tante le strategie che possono essere seguite per condurre disegni di ricerca che hanno l'algoritmo come oggetto di studio. In ognuno di questi, il ruolo di un ricercatore sociale di tecniche di ricerca sociale (sia quantitative che qualitative) è determinante per isolare gli assemblaggi socio-tecnici in cui avvengono le decisioni automatizzate, e stimarne gli impatti. Solo la ricerca empirica può consentire di mettere da parte il pregiudizio ed entrare invece nel merito del funzionamento di questi sistemi. Quante più ricerche si realizzeranno, tanto più sarà possibile analizzare il feedback continuo che c'è tra algoritmi, conoscenza prodotta con gli algoritmi e decisioni.

4. Cosa può fare la politica

L'esigenza di rendere trasparenti i sistemi di decisione algoritmica, e in generale tutte le tecnologie *data intensive*, è stato discusso a vari livelli di governo – come ad esempio alla Casa Bianca e al Parlamento Europeo – ma esistono ancora degli ostacoli alla diffusione dell'audit di algoritmi. Secondo un recente report del parlamento europeo *Panel for the Future of Science and Technology* ci sono degli ostacoli alla trasparenza che non si possono ignorare. In primo luogo, rendere pubblici i dati che gli algoritmi analizzano, o da cui «apprendono», può significare una violazione della privacy. Devono essere messe in campo delle procedure di anonimizzazione molto efficaci, per evitare che si verifichino infrazioni ai codici di riservatezza. Un altro limite è che queste tecnologie sono spesso protette da brevetti e diritti d'autore, per cui le operazioni e i codici alla base degli algoritmi non sono sempre accessibili (Geiger, 2017). Ma soprattutto, la trasparenza tramite l'apertura non porta molto lontano, perché comunque la comprensione del funzionamento degli algoritmi richiede competenze esperte, che non sono possedute dalla maggior parte dei cittadini, e spesso neanche dagli studiosi. La soluzione quindi non è un'apertura indiscriminata, ma piuttosto un'analisi esperta e indipendente, casomai affidata a degli organismi di sorveglianza indipendenti, riconosciuti e riconoscibili.

La costituzione di organismi terzi di controllo, che possano regolare l'impatto degli algoritmi e accertare se ci siano violazioni dei diritti, applicazioni scorrette e processi fallaci, consentirebbe di creare un clima di fiducia nei sistemi di decisione automatizzata, e un apparato di monitoraggio degli effetti dei sistemi di governo *data intensive*. Questi organismi dovrebbero coordinarsi con le autorità a protezione dei dati, con le altre autorità garanti (ad esempio l'Autorità per le Garanzie nelle Comunicazioni), e aderire a standard e norme di certificazione, come ad esempio quelle sviluppate dall'*International Standards Organization* (Iso). Sono molte le procedure organizzative che sono sottoposte a certificazione, perché non certificare anche le procedure algoritmiche, se queste possono generare disuguaglianze sociali?

Parallelamente alla fondazione di questi organismi di controllo, andrebbe avviata anche una seria campagna di educazione alla consapevolezza da parte di particolari gruppi sociali, e dei cittadini in genere, del potere che può essere esercitato tramite gli algoritmi. C'è la necessità di fornire una massiva alfabetizzazione algoritmica, insegnando concetti chiave

come il pensiero computazionale adottato per prendere decisioni, il *co-ding*, l'importanza dei dati e i loro usi, la difesa dei dati personali, ecc. La mancanza di questa formazione limita la capacità di azione degli individui nei sistemi algoritmici, con la conseguenza inattesa di indebolire la pressione sul mercato e sui governi a regolamentare il settore.

Infine, c'è anche l'esigenza di formare figure professionali riconosciute che siano capaci di studiare gli assemblaggi e valutare gli algoritmi. Gli auditor di algoritmi sono indispensabili per comprendere le conseguenze che possono avere sui risultati le interfacce delle piattaforme, le architetture delle infrastrutture dati e i modelli scelti. Si tratta di figure professionali molto ibride, caratterizzate da competenze interdisciplinari che intrecciano la ricerca sociale, l'informatica, la *data science*, l'etica e il diritto. Le grandi multinazionali di consulenza stanno lavorando proprio su come mettere in campo professionisti in grado di proporre soluzioni per il monitoraggio dei sistemi di decisione algoritmica e delle loro performance. Con la città di Amsterdam, ad esempio, Kpmg insieme alla studiosa americana O'Neil stanno conducendo analisi dei rischi e controlli ambientali per rilevare la capacità di organizzare effettivamente il controllo dell'adozione di soluzioni automatizzate. I loro studi riguardano non solo gli effetti sui cittadini-utenti, ma anche l'analisi degli impatti che questi sistemi hanno sulle organizzazioni stesse che li adottano. Queste soluzioni proattive di controllo, messe in campo da partenariati pubblici e privati sotto la sorveglianza di organismi terzi, che vedono un ruolo importante dei ricercatori, possono essere davvero le forme più efficaci per alimentare la fiducia verso i sistemi di decisione automatizzata, anticipandone i rischi.

5. Conclusioni

Il principale campo di applicazione degli algoritmi è la burocrazia (Visentin, 2018).

Con questo articolo si è voluto dimostrare che il successo dell'applicazione degli algoritmi nella pubblica amministrazione dipenderà dalla capacità di produrre evidenze empiriche riguardo alla loro qualità, al loro funzionamento nella società e ai rischi che hanno nel generare disuguaglianze. Un algoritmo non è buono o cattivo di per sé. Pertanto, la posizione che si dovrebbe assumere rispetto all'adozione di procedure algoritmiche nella gestione dei servizi pubblici è una posizione intermedia, che sia critica rispetto all'uso indiscriminato e incontrollato degli

algoritmi, ma che, allo stesso tempo, veda di buon occhio un impiego consapevole e trasparente delle decisioni automatizzate. Essere consapevoli non vuol dire ostacolare la diffusione di sistemi decisionali *data intensive*, ma, anzi, consente alla lunga di migliorare il loro operato.

Il caso delle applicazioni per il tracciamento dei movimenti della popolazione al fine di contenere la diffusione del contagio da Covid-19 può servire da esempio emblematico per concludere il percorso condotto in questo lavoro, e chiarire cosa significhi davvero adottare un approccio critico e consapevole ai sistemi di decisione automatizzata.

Nel dibattito pubblico internazionale è passato il messaggio che solo le applicazioni software più invasive per la vita privata dei cittadini possano controllare la propagazione del nuovo coronavirus. È quanto è stato detto in Corea del Sud, in Italia e negli USA, e in chissà quanti altri paesi, ma in realtà non ci sono evidenze empiriche a sostegno di questa convinzione. *Algorithm Watch*, organismo indipendente di ricerca sugli algoritmi, ha pubblicato un appello affinché si smetta di considerare sicurezza (privacy) ed efficacia delle tecnologie come alternative. Prima di analizzare le implicazioni che queste applicazioni hanno sulla privacy, dovremmo innanzitutto domandarci se servono davvero. I pochi risultati di ricerca che si trovano in letteratura sono molto disparati, e dipendono fortemente dal contesto (Ferretti e al., 2020). Inoltre, esistono delle applicazioni che si basano su approcci più trasparenti e garanti dei diritti rispetto a quelle adottate in Corea del Sud. L'applicazione *Safe Paths* sviluppata dal Mit, oppure l'applicazione *TraceTogether* utilizzata a Singapore, sono esempi del fatto che la protezione della privacy e la protezione dal coronavirus non siano mutuamente esclusive. Prima di adottare sistemi di tracciamento, pertanto, dovrebbero essere ricercate evidenze dei loro benefici, comunicando chiaramente ai cittadini chi li costruisce e li diffonde, specialmente se sono coinvolti soggetti privati. E, quando vengono dimostrati efficaci, i sistemi di sorveglianza digitale andrebbero implementati in osservanza di regolamenti per la tutela della protezione dei dati, come il regolamento europeo Gdpr. Come è stato evidenziato dal comitato Europeo per la Protezione dei Dati, i cittadini dovrebbero avere il diritto di ricorrere ad ogni decisione presa da un sistema automatizzato che non riconoscono come equa (ad esempio, che un individuo è stato a contatto con una persona infetta e deve rimanere in quarantena).

Rimane il grande limite che, come si è cercato di dimostrare in questo articolo, l'adozione di sistemi di decisione algoritmica può discriminare alcune categorie di soggetti. Nel caso del Covid-19, bisogna essere con-

sapevoli che un algoritmo che calcola il rischio sulla base di dati raccolti in modo anonimo e aggregato, potrebbe associare un rischio sistematicamente elevato a certe categorie di persone (ad esempio, addetti alle consegne alimentari, corrieri, tassisti, *runners*, ecc.) (Loi, 2020). E quando a una determinata categoria sociale è associata una probabilità di rischio molto alta, essa può venire stigmatizzata socialmente.

Per quanto sia difficile, è unicamente attraverso un attento esame dei pro e dei contro basato sui risultati di ricerca, che è possibile adottare consapevolmente un sistema di decisione automatizzato. L'audit di algoritmi è lo strumento cardine per vigilare sull'autorità algoritmica, altrimenti il rischio è che la governance algoritmica continui a guidare le nostre vite senza il nostro intervento, e a produrre nuove, e sempre più complesse, forme di disuguaglianza sociale da cui sarà difficile, e in certi casi impossibile, proteggersi.

Riferimenti bibliografici

- Airoldi M., 2020, *Lo spettro dell'algoritmo e le scienze sociali. Prospettive critiche su macchine intelligenti e automazione delle disuguaglianze*, «Polis», vol. 35, n. 1, pp. 111-128.
- Aragona B., 2020, *Come intervenire sulle decisioni automatizzate: l'audit di algoritmi*, intervento al XII Congresso nazionale AIS «Sociologia in dialogo: algoritmo, cervello, valutazione», 23 gennaio 2020, Università di Napoli Federico II, Napoli.
- Aragona B. e Felaco C., 2019, *Big Data from Below. Researching Data Assemblages*, «Tecnoscienza: Italian Journal of Science & Technology Studies», vol. 10, n. 1, pp. 51-70.
- Broussard M., 2018, *Artificial Unintelligence: How Computers Misunderstand the World*, Mit Press, Cambridge (Ma).
- Dourish P., 2016, *Algorithms and their Others: Algorithmic Culture in Context*, «Big Data & Society», vol. 3, n. 2, pp. 1-11.
- Eubanks V., 2017, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, St. Martin's Press, New York.
- Ferretti L., Wymant C., Kendall M., Zhao L., Nurtay A., Abeler-Dörner L. e Fraser C., 2020, *Quantifying Sars-CoV-2 Transmission Suggests Epidemic Control with Digital Contact Tracing*, «Science», 31 marzo, disponibile all'indirizzo internet: <https://science.sciencemag.org/content/early/2020/04/09/science.abb6936/tab-pdf>.
- Geiger R.S., 2017, *Beyond Opening up the Black Box: Investigating the Role of Algorithmic Systems in Wikipedian Organizational Culture*, «Big Data & Society», vol. 4, n. 2, pp. 1-14.

- Gillespie T., 2014, *The Relevance of Algorithms*, in Gillespie T., Boczkowski P. e Foot K.A. (a cura di), *Media Technologies: Essays on Communication, Materiality, and Society*, Mit Press, Cambridge (Ma), pp. 167-194.
- Halford S. e Savage M., 2010, *Reconceptualising Digital Social Inequality*, «Information, Communication and Society», vol. 13, n. 7, pp. 937-955.
- Hallinan B. e Striphas T., 2016, *Recommended for You: The Netflix Prize and the Production of Algorithmic Culture*, «New media & society», vol. 18, n. 1, pp. 117-137.
- Kayser-Bril N., 2020, *Google Apologizes after its Vision AI Produced Racist Results*, «Algorithm Watch», 7 aprile 2020.
- Kitchin R., 2017, *Thinking Critically about and Researching Algorithms*, «Information, Communication and Society», vol. 20, n. 1, pp. 14-29.
- Kitchin R. e Dodge M., 2011, *Code/Space: Software and Everyday Life*, Mit Press, Cambridge (Ma).
- Latour B., 1987, *Science in Action: How to Follow Scientists and Engineers through Society*, Harvard University Press, Cambridge (Ma).
- Loi M., 2020, *We Must Save Privacy from Privacy Itself*, «Algorithm Watch», 31 marzo 2020.
- Lupton D., 2015, *Digital sociology*, Routledge, Londra.
- Musiani F., 2013, *Governance by Algorithms*, «Internet Policy Review», vol. 2, n. 3, pp. 67-89.
- Noble S., 2018, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York University Press, New York.
- O'Neil C., 2016, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Random House, New York.
- Pasquale F., 2015, *The Black Box Society: The Secret Algorithms that Control Money and Information*, Harvard University Press, Cambridge (Ma).
- Rogers R., 2013, *Digital Methods*, Mit Press, Cambridge (Ma).
- Salvucci A., Giorgi M., Barchiesi E. e Scafidi M., 2017, *Perizia tecnica preliminare sull'analisi dell'algoritmo che gestisce il software della mobilità docenti per l'a.s. 2016/2017*.
- Sandvig C., Hamilton K., Karahalios K. e Langbort C., 2014, *Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms*, intervento presentato al convegno «Data and Discrimination: Converting Critical Concerns into Productive Inquiry» del 64th Annual Meeting of the International Communication Association, 22 maggio 2014, Seattle.
- Sipser M., 2006, *Introduction to the Theory of Computation*, Thompson Course Technology div. of Thompson Learning, Boston.
- Takhteyev Y., 2012, *Coding Places: Software Practice in a South American City*, Mit Press, Cambridge (MA).

Visentin C., 2018, *Il potere razionale degli algoritmi tra burocrazie e nuovi idealtipi*, «The Lab's Quarterly», vol. 20, n. 3, pp. 47-72.

Zunino C., 2019, *L'algoritmo impazzito fu contro la costituzione*, «La Repubblica», 17 settembre 2019.

RPS