RISCHI ALGORITMICI E STRUMENTI DI MITIGAZIONE

Biagio Aragona, Francesco Amato
Università degli Studi di Napoli Federico II
aragona@unina.it – francesco.amato2@unina.it

Abstract – Aziende e pubbliche amministrazioni usano con sempre più frequenza algoritmi per prendere decisioni sulle vite di clienti, cittadini, utenti e pazienti che incidono su, ad esempio, la scelta del prodotto da proporgli, la concessione di un prestito, l'assegnazione di un posto di lavoro, l'erogazione di una certa quantità di farmaco. I dati personali alimentano la macchina del capitalismo digitale, diventando moneta di scambio per l'accesso ai servizi online. Attraverso questa transazione informazionale, spesso celata agli utenti, gli algoritmi di profilazione estraggono valore e comportano potenziali conseguenze sociali per gli individui, come la sorveglianza digitale. Altri studi hanno dimostrato che i sistemi di decisione automatizzata possono generare, e soprattutto amplificare, le disuguaglianze sociali. La scarsa conoscenza del coding e della data science, così come l'accesso limitato alla tecnologia e a Internet, spesso comportano una scarsa capacità di comprendere i rischi del processo decisionale algoritmico. Come è possibile mitigare questi rischi? La ricerca sociale può intervenire su questi temi indirizzando la diffusione degli algoritmi nella società verso un futuro inclusivo riflessivo e critico.

Parole chiave: Rischio algoritmico; Sorveglianza digitale; Algorithm audit; Disuguaglianze socio-digitali; Privacy.

1. Introduzione

Da tempo viviamo in una società in cui i dati e le tecnologie digitali svolgono progressivamente un ruolo sempre maggiore. Dal credito alla salute, dall'istruzione al lavoro, sono sempre più gli ambiti sociali che sono regolati da algoritmi che suggeriscono, o prendono decisioni, su numerosi aspetti della nostra esistenza. Auto a guida autonoma, robot che automatizzano il lavoro di cura e di assistenza, tecnologie che abilitano nuove forme di apprendimento e di organizzazione del lavoro, dispositivi per il monitoraggio in tempo reale delle condizioni di salute, sensori che governano i sistemi dei trasporti locali, avatar che si relazionano all'interno del metaverso, sono solo alcuni degli esempi che testimoniano quanto oramai gli algoritmi, e i dati che elaborano, siano entrati pervasivamente a far parte della nostra vita sociale. Tutta questa innovazione tecnologica ha senza dubbio generato nuove opportunità, ma, al contempo, ha portato alla luce nuove sfide per la società, e nuovi rischi, tra cui il rischio algoritmico.

Per rischio algoritmico si intende ogni conseguenza dannosa intenzionale o non intenzionale che gli algoritmi possono avere su individui o gruppi sociali specifici. Un primo esempio di rischio algoritmico può essere generato dai sistemi che suggeriscono, o prendono decisioni in specifici contesti istituzionali. Ad esempio, l'adozione di questi sistemi per la selezione della platea dei beneficiari delle misure di welfare ha avuto in alcuni casi conseguenze pesanti sulle vite dei cittadini più poveri e socialmente esclusi. Una nota studiosa afroamericana ha sottolineato, riguardo all'assegnazione dei posti letto ai senza fissa dimora della città di Los Angeles, che l'algoritmo per il calcolo dell'indice di vulnerabilità andava a contare le notti passate in prigione come housing, abbassando di fatto il valore dell'indice a chi veniva arrestato, riducendo così le possibilità di questa persona di accedere in futuro ai pochi posti letto disponibili (Eubanks, 2018). In Italia, invece, i sistemi di decisione algoritmica 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. In circa 10.000 casi l'algoritmo prese in considerazione unicamente le preferenze espresse dai candidati, senza fare un confronto tra i punteggi e le destinazioni.

ISSN 2724-1971

Il risultato fu che inviò insegnanti pugliesi e docenti di Catanzaro in provincia di Milano, quando avrebbero dovuto essere destinati alle loro regioni e, in maniera del tutto incomprensibile, spedì a Prato due professori calabresi con i figli autistici. In modo diverso anche gli ecosistemi digitali e i loro algoritmi che processano dati possono produrre nuovi rischi. Ad esempio, se da un lato il Metaverso di Facebook offre agli utenti opportunità per lavorare, giocare, connettersi con gli altri, dall'altro è gia stato tristemente riportato il primo caso di violenza sessuale di gruppo all'interno dell'ecosistema. Una psicoterapeuta quarantenne ha dichiarato di essere stata accerchiata da diversi avatar maschili e di aver subito una vera e propria violenza di gruppo. Seppure si tratti di un evento avvenuto in un ambiente digitale, le conseguenze psicologiche sulla vittima non sono state trascurabili, sottolineando la necessità di definire strumenti che siano in grado di anticipare i rischi degli algoritmi e degli ecosistemi digitali, e di prevenire l'accadere di discriminazioni, ingiustizie e abusi, mediati e automatizzati dalle tecnologie.

2. La percezione del rischio algoritmico

Alla costante diffusione degli algoritmi nella società corrisponde solo una limitata conoscenza del loro funzionamento da parte della popolazione, e, quindi, degli effetti che questi possono generare sulle persone e sulla società. Individui con scarse competenze digitali e ridotta capacità di accesso alla rete e alla tecnologia tendono a sottostimare i rischi degli algoritmi e, allo stesso tempo, ad avere poca fiducia nei sistemi automatizzati. Ad esempio, la popolazione più anziana e meno digitalizzata ha maggiori difficoltà a comprendere i meccanismi sottesi all'utilizzo di un algoritmo, mentre le nuove generazioni, cresciute sulle piattaforme della società digitale, potrebbero riporre molta fiducia in questi strumenti. Il rischio algoritmico è, quindi, un fenomeno che non è percepito in modo uniforme, anche perché è l'opacità stessa degli

algoritmi a renderlo sfuggente.
Il funzionamento degli algoritmi non è direttamente conoscibile, perché è racchiuso nella "scatola nera", ed è possibile accedere solo ai dati iniziali e ai risultati, ma non all'intero processo. Aprire la scatola nera vuol dire comprendere gli elementi che costituiscono l'algoritmo: gli apparati, le tecniche, e le diverse comunità di attori che hanno partecipato al suo assemblaggio. Infatti, malgrado si sottolinei sempre l'aspetto tecnico degli algoritmi, questi sono invece il risultato di un'azione umana. Interessi privati e pubblici sono il fattore essenziale che spinge verso la realizzazione di algoritmi, al fine di automatizzare e rendere possibili determinati processi.
Considerare gli algoritmi come il frutto della cultura umana e degli interessi di numerosi attori, evidenzia come il codice non sia un prodotto esclusivamente tecnico, ma che contiene, al proprio interno, tracce delle influenze socio-culturali degli individui che hanno portato alla sua realizzazione. Pertanto, quando parliamo di algoritmi, sarebbe più corretto fare riferimento ad un costrutto socio-tecnico, proprio per evidenziare la dualità sociale e tecnologica insita nella produzione algoritmica. Comprendere la socio-tecnicità degli algoritmi è un primo passo fondamentale per identificare come essi possano replicare e rinforzare stereotipi culturali della nostra società. Ad esempio, è stato notato che i motori di ricerca tendono a perpetuare gli stereotipi di genere e una cultura maschilista, che da un lato rinforza l'immagine della donna come oggetto di desiderio sessuale e dall'altro la relega sempre ai ruoli tradizionali di mamma e lavoratrice domestica (Noble, 2018). Gli effetti che, quindi, possono essere attuati dagli algoritmi riguardano fenomeni quali la discriminazione sociale, di genere e di razza e l'esclusione dall'acquisto e dall'utilizzo di prodotti in base a differenze di reddito, culturali e sessuali. Gli effetti degli algoritmi non impattano, poi, solo sul singolo utente di internet, ma anche su specifiche categorie di cittadini sempre più spesso costrette ad utilizzare i dispositivi digitali e la rete per usufruire di servizi fondamentali, quali la propria identità digitale, la richiesta di certificati e documenti, e le operazioni finanziarie. Sotto questa prospettiva, il rischio algoritmico può rappresentare un'importante sfida sia per i diritti fondamentali dell'individuo, come la privacy, la libertà e l'uguaglianza, ma anche per i diritti e i valori collettivi, tra cui l'equità, la sicurezza, l'inclusività, la responsabilità e il controllo democratico (Aragona, 2021).

3. Tipi di rischio algoritmico

Le finalità degli algoritmi rispondono a numerosi obiettivi, uno dei principali è sicuramente la profilazione automatizzata, attraverso cui gli algoritmi consentono di ottenere le informazioni personali degli utenti, come, ad esempio, il dispositivo utilizzato, la località di accesso e gli usi della rete internet. La raccolta automatizzata dei dati è un'azione algoritmica molto controversa, poiché interessa direttamente la privacy degli individui, consentendo l'acquisizione delle informazioni personali senza che gli utenti abbiano la percezione del processo. Attraverso l'incremento della consapevolezza digitale, garantita a tutte le fasce della popolazione, sarebbe possibile fornire ai cittadini i primi strumenti di difesa per ottenere una migliore comprensione dei rischi algoritmici legati alla propria privacy. La violazione della riservatezza dei propri dati, infatti, non va affrontata esclusivamente attraverso l'incremento delle competenze digitali della cittadinanza, ma anche garantendo la sicurezza delle informazioni degli utenti e dei cittadini tramite il ricorso ad autorità specifiche, capaci di prevenire e mitigare ulteriori effetti degli algoritmi, come, ad esempio, la sorveglianza digitale, le pratiche di Redlining e le disuguaglianze sociali digitali.

Privacy

È evidente come la privacy rappresenti una delle principali questioni toccate dal rischio algoritmico, in quanto è con la rielaborazione e l'utilizzo di dati che gli algoritmi della società digitale producono valore. Il primo strumento a sostegno degli utenti è proprio il diritto alla privacy che, sempre più spesso riconosciuto dalle normative per la protezione dei dati personali, garantisce la piena tutela delle informazioni sensibili delle persone, e consente loro una maggiore libertà nello svolgimento delle proprie attività essenziali, senza il rischio di dover essere sottoposti ai trattamenti di profilazione non autorizzati (Tsamados, et al., 2022).
Il diritto alla privacy, soprattutto negli anni passati, è stato spesso aggirato dai sistemi algoritmici attraverso la subdola acquisizione del consenso degli utenti ai trattamenti. Una volta che gli utenti cedono le proprie informazioni personali, difficilmente riescono ad avere piena consapevolezza delle analisi, degli usi e delle finalità alla quale sono sottoposti i propri dati. Ciò accade poiché i moderni sistemi algoritmici, rendendo silente questo processo, da un lato, non notificano nulla all'utente, dall'altro, invece, precludono persino la possibilità di verificare quale trattamento dei dati sia davvero avvenuto.

Sicurezza

I dati personali, raccolti e rielaborati dalle imprese digitali, sono elementi fondamentali delle nostre società

digitali, tanto da essere diventati in breve tempo la principale valuta dell'economia in rete. Queste informazioni vengono cedute dagli utenti quando accettano i trattamenti sui dati personali presenti sui servizi e sui prodotti digitali. Alcune compagnie di comunicazione e logistica sono state sanzionate per aver fatto un uso improprio dei dati dei propri utenti. Ad esempio, Facebook è stato multato per pratica commerciale ingannevole, perché

al momento dell'attivazione dell'account, informava gli utenti unicamente della gratuità dell'iscrizione, mentre ometteva che i dati personali, sarebbero stati ceduti al social media e utilizzati per scopi commerciali.

> La cessione delle informazioni non sempre è chiaramente avvertita dall'utente. Infatti, Bauman (2015) afferma che la cessione di porzioni delle proprie identità e privacy per l'accesso ai prodotti della rete è sempre più spesso considerata un costo ragionevole, in quanto, gli utenti non hanno gli strumenti adeguati per valutare le conseguenze delle proprie azioni.

Questo fenomeno, unito alla percezione della gratuità del prodotto, può indurre erroneamente l'utente a pensare di accedere ai contenuti in modo completamente gratuito quando, poi, invece, pagano attraverso la cessione delle proprie informazioni, dallequali possono svilupparsi gravose conseguenze sociali come, ad esempio, quella della sorveglianza digitale.

Sorveglianza digitale

La sorveglianza digitale, si configura come il controllo sistematico dei dati personali effettuato sulla popolazione attraverso il continuo monitoraggio dei dispositivi e degli usi della rete internet. Esempi in tal senso ce ne sono già, anche in ambito pubblico, basti pensare alle attività di sorveglianza di massa messe in campo dall'Agenzia per la Sicurezza Nazionale statunitense (NSA). L'NSA ha collezionato metadati sulle telefonate effettuate attraverso tutti i gestori statunitensi e una speciale divisione dell'agenzia, chiamata Follow the Money, raccoglieva dati sulle transazioni finanziarie di importanti istituti internazionali come Visa, Mastercard e SWIFT. Attraverso il programma di sorveglianza PRISM, l'NSA ha avuto accesso diretto ai server di molte delle principali aziende dell'informatica statunitense, quali Microsoft, Google, Yahoo!, Facebook, Apple, YouTube e Skype.

L'agenzia monitorava quindi le attività degli utenti, compresi scambi di messaggi, foto e video, conservando in particolare le liste di indirizzi degli utenti usate nei servizi e-mail e di messaggistica istantanea. Il trattamento di questa immensa mole di dati veniva giustificata per questioni di sicurezza e legata alla lotta al terrorismo e alla salvaguardia dello stato.

Profilazione

Gli utenti vengono attribuiti a determinati profili che rappresentano la popolazione, come una moderna configurazione delle classi sociali. Queste suddivisioni in gruppi possono, a volte, condurre alla sistematica esclusione di determinate porzioni di popolazione da prodotti e servizi. In questo modo il rischio di incorrere in importanti fenomeni di disuguaglianza è molto concreto. Di conseguenza, i contenuti mostrati sui motori che verrebbero mostrati sui motori di ricerca o sui social network non sarebbero i medesimi per gruppi di persone appartenenti a gruppi diversi. Gli utenti, spesso inconsapevoli degli effetti degli algoritmi, utilizzerebbero la rete in sostanziali condizioni di asimmetria informativa, incapaci di attuare manovre di salvaguardia e di mitigazione dei propri rischi.

Redlining

Simili fenomeni possono generare pericolosi effetti discriminatori, come il Redlining, ovvero la pratica attraverso cui prodotti e servizi vengono sottratti a potenziali clienti classificati come pericolosi o, semplicemente, sgraditi. Il redlining digitale rievoca la pratica storica risalente agli anni '30 e avvenuta negli Stati Uniti e in Canada dove, venivano tracciate linee rosse sulle mappe per indicare i poveri e principali quartieri neri ritenuti inadatti a prestiti finanziari, che hanno contribuito a accrescere le già grandi disparità economiche tra i quartieri. Ecco che la cessione delle proprie informazioni, raccolte dagli algoritmi, apre le porte a potenziali rischi che potrebbero impattare – negativamente – sulla vita personale dei cittadini e non limitarsi alla comparsa di pubblicità fastidiose durante una sessione di navigazione online.

Disuguaglianze sociali digitali

Infine, la scarsa trasparenza del processo algoritmico, nonché la difficile attribuzione delle responsabilità dello stesso, evidenziano ulteriori fattori di rischio. Attualmente, senza essere un esperto informatico, è molto complesso avere piena consapevolezza del codice e delle sue specificità. La costruzione degli algoritmi è un processo eterogeneo, al quale prendono parte numerosi attori che, con le proprie influenze socio-culturali, partecipano alla sua realizzazione.

> ***"Alla costante diffusione degli algoritmi nella società corrisponde solo una limitata conoscenza del loro funzionamento da parte della popolazione, e, quindi, degli effetti che questi possono generare sulle persone e sulla società."***

Per queste ragioni, il tema dell'attribuzione della responsabilità è quantomeno complesso. Sebbene in Europa la legislazione preveda che gli individui abbiano il diritto di non essere sottoposti a processi decisionali automatizzati, spesso i decisori si limitano, però, a confermare l'istruzione ottenuta dall'elaborazione algoritmica senza discuterne l'affidabilità. Gli effetti di queste azioni possono essere molto controversi, in quanto la scelta finale è stata sì presa dall'agente umano, ma tale decisione è frutto del risultato di un processo decisionale automatizzato e non controllato appieno..

4. Proposte di mitigazione del rischio algoritmico: tra politica e ricerca

La mitigazione dei rischi algoritmici è un'azione che, per essere messa in campo, richiede innanzitutto la comprensione da parte di tutti gli stakeholder, cittadini e utenti in primis, di questo tipo di rischio. Allo stato attuale non esiste un'unica ricetta, anche perché lo stesso studio degli effetti degli algoritmi è ancora un processo in divenire.

Non neutralità algoritmica

L'accettazione degli algoritmi e, in generale, delle piattaforme digitali come strumenti non neutrali è, sicuramente, un primo importante passo che dovrebbe emergere anche al di

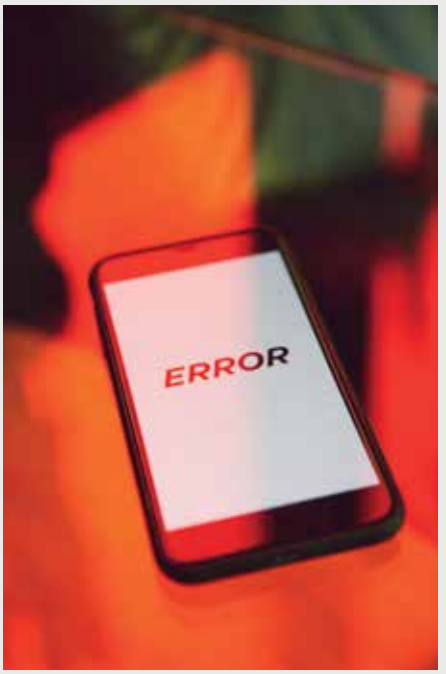
fuori della comunità scientifica, ed essere ricompreso entro gli obiettivi delle agende politiche. Su queste premesse si sono realizzati numerosi studi che contestano con forza che gli algoritmi, e i dati su cui vengono eseguiti, siano strumenti obiettivi, imparziali e affidabili by design, ribadendo 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.

Incoraggiare la ricerca sociale sugli algoritmi
Inoltre, è necessario sottolineare come il metodo della ricerca sociale si è rivelato, in questi primi studi sul fenomeno, un valido strumento per l'indagine algoritmica. Riteniamo che alcune tecniche della ricerca sociale, sia quantitative che qualitative, condotte on-line o off-line, come, ad esempio, l'etnografia, gli esperimenti, e le survey, apporteranno sempre di più importanti contribuiti alla comprensione del funzionamento degli algoritmi e all'analisi dei loro impatti sulla società.

Istituire organismi terzi di controllo
La costituzione di organismi terzi di controllo, che possano regolare l'impatto delle tecnologie ad uso intensivo di dati 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 algoritmici di governo. Questi organismi dovrebbero coordinarsi con le autorità a protezione dei dati e con le altre autorità garanti (ad esempio l'Autorità per le Garanzie nelle Comunicazioni).

Il registro pubblico dei sistemi algoritmici
Attraverso l'istituzione di un registro pubblico degli algoritmi, sarebbe garantita una maggiore trasparenza e responsabilità, consentendo ai cittadini e a tutti gli attori interessati di essere più sicuri non solo riguardo all'uso dei servizi digitali, ma anche dei processi automatizzati che vengono attivati a loro insaputa sulle rispettive informazioni personali.

Educare alla consapevolezza algoritmica
Parallelamente alla fondazione del registro pubblico e all'istituzione 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 necessita di fornire una massiva alfabetizzazione algoritmica, insegnando concetti chiave come il pensiero computazionale adottato per prendere decisioni, il coding, 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.

Formare auditor di algoritmi
C'è inoltre l'esigenza di formare figure professionali riconosciute che siano capaci di 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 computazionali 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.

Algorithm Audit
Infine, la maggiore comprensione del rischio algoritmico sarà facilitata dall'identificazione di un protocollo, sviluppato attraverso il coinvolgimento dei decisori politici, delle università, delle imprese, della cittadinanza e di tutti gli attori pubblici e privati interessati. Il protocollo dovrà essere sviluppato a partire da caratteristiche e criteri che rispondano ai requisiti di trasparenza e di responsabilità già validati e impiegati in altri settori economici e sociali.

Bibliografia

Aragona, B. (2021). *Algorithm Audit: Why, What, and How?* (1st ed.). Routledge

Bauman, Z. e Lyon, D. (2015). *Sesto potere. La sorveglianza nella modernità liquida*. Laterza

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press

Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism.* New York University Press

Tsamados, A., Aggarwal, N., Cowls, J. et al. (2022). *The ethics of algorithms: key problems and solutions.* AI & Soc 37, 215–230