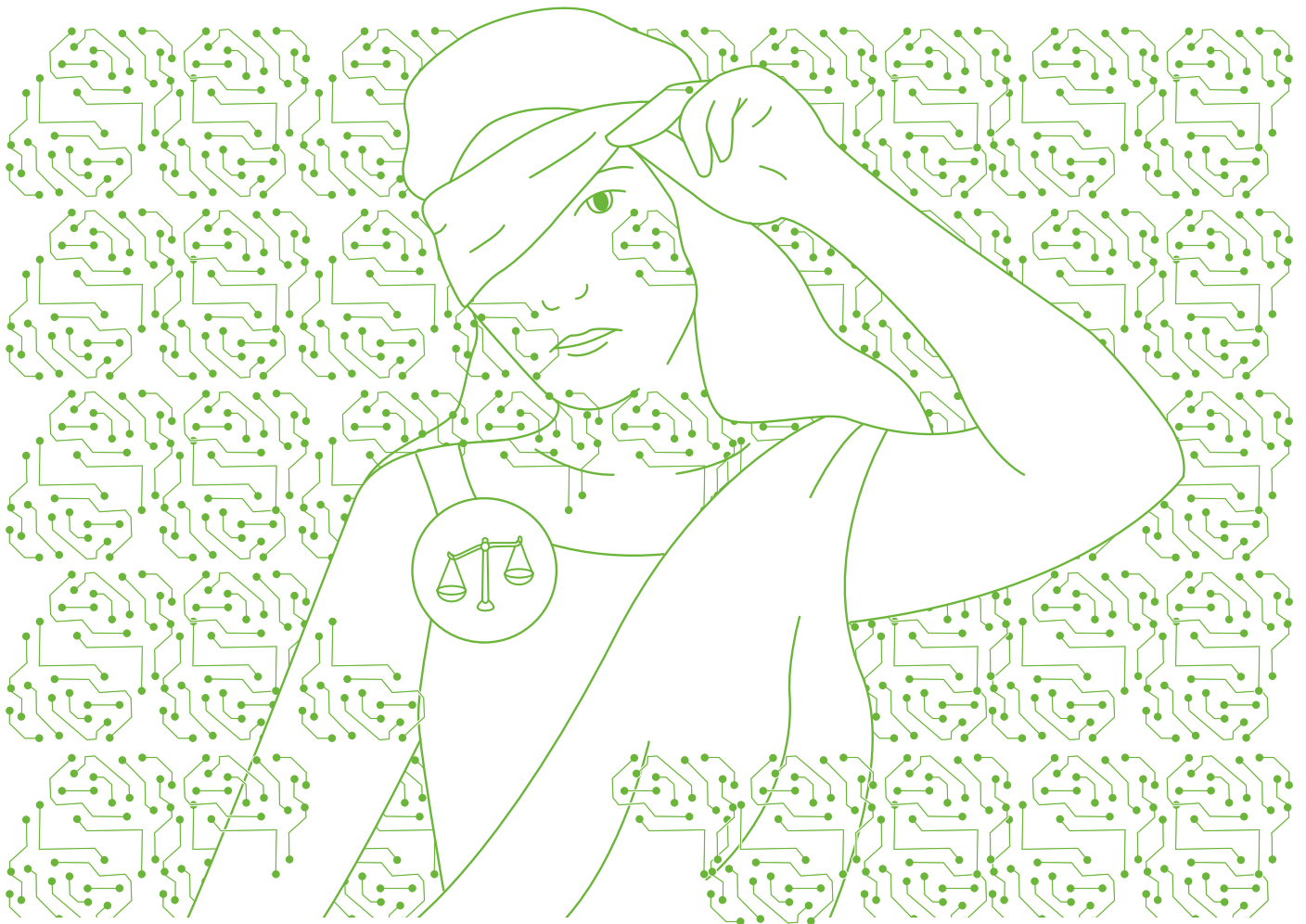


L'intelligenza artificiale non è più discriminante dei dati che le vengono forniti

Intervista a **Henrike Schlottmann**
Jan Kuhlen



L'intelligenza artificiale (IA), sempre più impiegata per i processi decisionali basati su algoritmi, si diffonde in tutti i rami e con essa anche il rischio di discriminazione di singoli individui e gruppi di persone. Henrike Schlottmann e Jan Kuhlen del «Zentrum für digitalen Fortschritt» (centro del progresso digitale) spiegano nell'intervista come ridurre detta disparità di trattamento. Loro sostengono la tesi secondo cui l'intelligenza artificiale sia soltanto lo specchio delle distorsioni sociali. Ritengono infatti che con la sistematizzazione degli errori sarà più facile correggere la discriminazione, sia a livello digitale che analogo.

Perché l'intelligenza artificiale è discriminante?

Jan Kuhlen: Con discriminazione s'intende il trattamento equo di disparità o la disparità di trattamento di cose o persone uguali, senza giustificazione alcuna. Sostanzialmente la tecnologia non è di per sé discriminante, tranne se gli esseri umani ne traggono conclusioni discriminanti o la istruiscono in modo tale da comportare automaticamente risultati discriminanti. Stesso dicasi per l'intelligenza artificiale. Il suo impiego può comportare una discriminazione involontaria perché non è chiaro come si ottengono risultati da una rete neurale artificiale. Al momento non si è consapevoli dei criteri determinanti per valutare bene le sfide legate all'intelligenza artificiale.

Quali forme di discriminazione sono note nell'ambito dell'intelligenza artificiale?

Henrike Schlottmann: La discriminazione può verificarsi su quattro livelli diversi: dipende dai dati del training, dagli algoritmi prescelti, dal team e dagli obiettivi prefissati. È la persona a optare per obiettivi discriminanti o a impiegare team che presentano già punti ciechi. Di norma, anche nell'impiego dell'IA, la discriminazione si basa sugli attributi protetti come sesso, colore della pelle o religione. In pratica si tratta della stessa problematica che s'incontra nella vita reale, poiché l'algoritmo alimentato tende soltanto a rispecchiare la distorsione, detta anche bias, che però esiste già nella società.

Potrebbe farci un esempio?

Henrike Schlottmann: Un esempio classico sarebbe l'algoritmo per il reclutamento presso Amazon¹ rivelatosi discriminante nel processo di candidatura: è infatti giunto alla conclusione che le candidature di soggetti femminili erano da considerare meno buone. Questo perché l'algoritmo è stato alimentato con un set di dati non equilibrato, in base a candidature pervenute ad Amazon negli ultimi dieci anni, la maggior parte delle quali erano state presentate da uomini. Il sistema ha poi dedotto da questo modello che si dovevano preferire i soggetti maschili. Un altro esempio è il «Predictive Policing», ossia la capacità predittiva dei crimini, secondo cui la persecuzione penale avviene in base a decisioni automatizzate. Negli USA sono così state fatte previsioni in merito alla recidività con un bias intrinseco nei dati di training che ha discriminato sistematicamente le persone «non bianche» in base al colore della pelle².

Per «discriminazione razionale» intende anche l'ottimizzazione dell'individuo a spese della solidarietà proprio come il sociologo Stefan Selke?

Jan Kuhlen: Con il «Lifelogging», vale a dire la raccolta di grosse quantità di dati relativi all'esistenza umana, la misurazione e la valutazione dell'individuo vengono intensificate. Questo potrebbe comportare anche una misurazione più stretta della società. Gli esseri

1: Künstliche Intelligenz diskriminiert (noch) [L'intelligenza artificiale è (ancora) discriminante]: <https://www.zeit.de/arbeit/2018-10/bewer-bungsroboter-kuenstliche-intelligenz-amazon-frauen-diskriminierung>

2: Predictive policing: Dem Verbrechen der Zukunft auf der Spur (Predictive policing: sulle tracce del crimine del futuro): <https://www.bpb.de/dialog/netzdebatte/238995/predictive-policing-dem-verbrecen-der-zukunft-auf-der-spur?type=galerie&show=image&i=238997>

umani possono così essere sempre più paragonati tra loro. Secondo i critici, detta misurazione intacca la nostra privacy.

Henrike Schlottmann: Il concetto di «razionalità» cela sempre anche qualcosa di sistematico. Per quanto riguarda la discriminazione questo significa che la deviazione sussiste già nella società e viene rispecchiata dall'algoritmo. La sua applicazione sistematizza ulteriormente il tutto. Non è quindi più l'essere umano a decidere in base alla situazione, bensì il sistema ad agire sistematicamente. Nel contempo, la discriminazione diventa più visibile e correggibile.

Vi sono ambiti di applicazione con una discriminazione razionale particolarmente marcata?

Jan Kuhlen: Nei dati sulla salute raccolti dalle app il rischio di discriminazione razionale è particolarmente elevato. Non si può invece dire lo stesso della discriminazione risultante dall'impiego dell'intelligenza artificiale. In questo caso il rischio di discriminazione sussiste soprattutto quando si tratta di valutare le persone, in particolare in merito agli attributi protetti.

Qual è il rapporto tra il «Lifelogging» e l'impiego d'intelligenza artificiale?

Henrike Schlottmann: «Lifelogging» o Big Data in sé non sono da considerare intelligenza artificiale, costituiscono tuttavia il presupposto per analisi intelligenti.

Jan Kuhlen: L'intelligenza artificiale, se la si impiega in questo modo, può effettuare in base a detti dati derivazioni e generalizzazioni che non sono improntate alla causalità. Ad esempio: chi si alza tutte le mattine alle 7 per correre tre chilometri ha più successo.

La discriminazione razionale esiste anche nella vita reale?

Jan Kuhlen: Sì. Se la compagnia assicurativa mi manda un questionario, l'analisi del rischio avviene anche in base ai dati così raccolti. Grazie a Big Data le valutazioni sono più dettagliate e possono essere sistematizzate più facilmente con l'intelligenza artificiale producendo risultati migliori: permettono di considerare in modo adeguato la situazione di rischio individuale. L'intelligenza artificiale sta facendo luce sull'intero sistema. Infatti la discriminazione razionale finora non è stata percepita abbastanza.

In un documento di sintesi³ chiedete regole vincolanti per l'impiego dell'IA al fine di evitare abusi.

Henrike Schlottmann: Chiediamo un consiglio etico globale il quale garantisca che determinate regole e linee guida valgano per tutti.

E questo come consentirebbe di evitare la discriminazione?

Jan Kuhlen: Sostanzialmente si deve tentare di evitare la discriminazione sui quattro livelli menzionati poc'anzi: il team deve essere selezionato in modo tale da evitare punti ciechi, i dati devono essere suddivisi in dati di apprendimento e di verifica e, in caso di dubbio, essere controllati con analisi di black box, ossia una procedura indiretta in grado di produrre a posteriori modelli di spiegazioni ad esempio attraverso visualizzazioni. Anche se la meccanica di una black box algoritmica, vale a dire di un sistema chiuso, non è nota, è possibile verificare il comportamento, quindi le immissioni ed emissioni.

Henrike Schlottmann: I sistemi devono essere possibilmente comprensibili e spiegabili. Si dovrebbe poter capire come sono state prese le decisioni. Gli algoritmi potrebbero ad esempio essere sottoposti a un audit; i dati sarebbero poi analizzati in base a bias. Gli stru-

3: Der Einfluss Künstlicher Intelligenz auf Freiheit, Gerechtigkeit und Solidarität (L'influsso dell'intelligenza artificiale su libertà, giustizia e solidarietà): <https://d-64.org/wp-content/uploads/2018/11/D64-Grundwerte-KI.pdf>

menti necessari a tale scopo dipendono tuttavia sempre dalla definizione personale di equità e sono lungi dall'essere ottimali.

Ma dobbiamo sempre poter argomentare con causalità per poter giustificare in modo plausibile le decisioni?

Jan Kuhlen: Per le decisioni statali a cui non possiamo sottrarci ci aspettiamo argomentazioni causali trasparenti e comprensibili. Simili decisioni non dovrebbero essere prese in base a una probabilità. L'intelligenza artificiale si basa tuttavia sempre su valori di probabilità. Ecco perché bisogna impiegare l'IA con prudenza. Non possiamo tuttavia aspettarci da una tecnologia che da un'immissione risulti sempre un'emissione del tutto comprensibile. Del resto non ce lo aspettiamo neanche dagli esseri viventi, tanto meno dall'uomo. La società dovrà abituarsi al fatto che in ambienti di laboratorio la tecnologia non produrrà mai risultati del tutto causali. Nel complesso può comunque essere molto più precisa dell'essere umano.

In quali ambiti è quindi legittimo avvalersi dell'IA?

Jan Kuhlen: Già oggi esistono numerose tecnologie con componenti di apprendimento automatizzate che funzionano bene: le traduzioni, le analisi video o il riconoscimento vocale. La tracciabilità, ossia il modo in cui un algoritmo ottiene buoni risultati di traduzione, interessa soltanto pochi. Alla fin fine, l'importante è che la frase sia ben tradotta. Anche nell'ambito del riconoscimento vocale - ad esempio la trasformazione di immissioni vocali in testo - per la maggior parte degli utenti è irrilevante sapere perché il sistema riconosce quanto detto.

Ma per quale motivo non si può allora stabilire che le tecnologie non comprensibili dovrebbero essere completamente proibite, perché potrebbero risultare sempre ingiustificatamente discriminanti?

Jan Kuhlen: Se la domanda si applicherebbe alle persone, dovremmo mettere in discussione anche le decisioni delle persone perché potrebbero risultare sempre ingiustificatamente discriminanti. Sostanzialmente l'intelligenza artificiale rappresenta una conquista con un enorme potenziale. Ma come già detto: in molti ambiti, ad esempio dove si devono prendere decisioni statali, occorre analizzarne più attentamente l'impiego.

L'introduzione di un consiglio etico globale suona illusorio poiché i principi normativi non valgono in tutto il mondo, com'è stato del resto ampiamente dimostrato dal gioco online «Moral Machine» del Massachusetts Institute of Technology (MIT)⁴ : i soggetti sono stati confrontati con situazioni di incidenti con auto semoventi e utenti della strada variabili e dovevano decidere quale fosse a loro avviso il danno più accettabile. Ne è risultato che i principi etici variano molto a seconda dell'area geografica.

Jan Kuhlen: Determinate regole e norme si sono imposte a livello mondiale. Anche nell'ambito dell'intelligenza artificiale non lo escludo. Il bando di sistemi d'arma autonomi viene ad esempio chiesto già nell'accordo di coalizione tra i partiti politici tedeschi CDU, CSU ed SPD⁵. È evidente che questo può funzionare soltanto a livello mondiale, come per le armi atomiche. Allo stesso modo i Paesi dovrebbero rendersi conto che è indispensabile segnalare che bot e avatar non sono esseri umani bensì macchine. Bisognerebbe inoltre chiarire

4: Moral Machine: Kann ein selbstfahrendes Auto ethisch handeln? (Moral machine: può un'auto semovente agire in modo etico?): <https://www.dw.com/de/moral-machine-kann-ein-selbstfahrendes-auto-ethisch-handeln/a-46045294-0>

5: Accordo di coalizione tra CDU, CSU e SPD: <https://www.bundesregierung.de/breg-de/themen/koalitionsvertrag-zwischen-cdu-csu-und-spd-195906>

quando le decisioni debbano essere spiegabili e comprensibili. Naturalmente è sempre più difficile trovare un codice di valori etico comune se si vogliono coinvolgere tanti Paesi.

Cosa accade se non ci riusciamo?

Henrike Schlottmann: Se osserviamo il cosiddetto Social Credit System possiamo intuire come sarebbe il mondo se gli USA o la Cina fissassero le proprie regole. Nel Regno di Mezzo si cerca attualmente di potenziare una tecnica di monitoraggio in grado di identificare immediatamente ogni singolo individuo in una grande massa. Chi si comporta bene viene premiato; per comportamenti negativi vi sono invece sanzioni. La domanda da porsi è quindi se, a lungo termine, questo è il mondo in cui desideriamo vivere? Trovo che dovremmo ambire a discutere determinate regole a livello internazionale perché l'intelligenza artificiale è un tema globale. Le sue applicazioni sono disponibili in tutto il mondo e grazie all'Internet tutti possono accedere a questo mercato.

Chiedete anche una gestione solidale dei dati. Cosa intendete esattamente?

Jan Kuhlen: Si tratta di ottenere un valore aggiunto ragionevole dai dati. Significa che i dati impersonali della società dovrebbero essere messi a disposizione in modo solidale per poter ottenere innovazioni sociali orientate al bene comune. Le semplificazioni ottenute grazie all'intelligenza artificiale potrebbero quindi essere utili anche per il servizio pubblico. Oggi molte idee naufragano per colpa di dati mancanti o della predisposizione inesistente degli stessi che richiede enormi risorse in termini di tempo. L'economia tedesca fa attualmente fatica a ottenere dati di qualità elevata. Al tempo stesso, i giganti della tecnologia statunitensi come Facebook e Google detengono molte informazioni in merito al comportamento dei consumatori tedeschi. Nell'economia tedesca si dispone anche di dati che però non vengono condivisi e sono troppo specifici.

Cosa proponete?

Jan Kuhlen: Si potrebbero ad esempio creare piattaforme dati intelligenti in grado di addestrare o scambiare dati. A tale scopo servono determinati standard che consentano agli attori di caricare i propri dati su una piattaforma per poi metterli a disposizione degli altri, gratuitamente o meno.

Henrike Schlottmann: Bisogna naturalmente accertarsi che i dati non siano personalizzati altrimenti potrebbero essere attribuiti a determinate persone. E non tutti vogliono che le informazioni sulla propria persona siano liberamente disponibili. Anonimizzare i dati è una procedura difficile perché non basta eliminare soltanto il nome. Si deve evitare che dagli altri attributi quali l'indirizzo o la formazione si possa desumere di quale persona si tratta.

Chiedete anche una gestione solidale dei dati sanitari?

Jan Kuhlen: Ritengo che la raccolta e la condivisione di dati sulla salute siano problematiche perché riguardano informazioni strettamente private. Si tratta niente meno che dello stato di salute di una persona. Non per nulla esiste il segreto medico. È risaputo che dette informazioni non devono essere condivise. Sostanzialmente non si tratta però della condivisione di dati personali, bensì di dati anonimizzati. Nel contempo riteniamo importante che i sistemi di dati siano protetti altrimenti avremmo un problema con l'autodeterminazione personale. A tutti serve la stessa protezione sanitaria, indipendentemente dal fatto che si vada a correre alle 7 di mattina o meno. Le casse malati non possono operare distinzioni basate su algoritmi, sia per quanto riguarda la gamma di prestazioni sia per quanto riguarda i costi.

La solidarietà è un concetto umano. Questo cosa significa per lo sviluppo di sistemi IA?

Jan Kuhlen: Sostanzialmente la tecnica è neutrale, ma ovviamente può essere utilizzata anche in modo solidale. Se grazie all'IA si riesce a gestire meglio le reti elettriche e a ridurre

di conseguenza il consumo energetico, ne approfitterebbero tutti. Tutte le utilizzazioni che si orientano al bene comune possono e devono essere promosse.

Henrike Schlottmann: Alla fin fine si tratta sempre anche di decisioni politiche. È evidente però che se il mio obiettivo non è solidale, non sarà solidale neanche il risultato. Se un algoritmo viene ad esempio utilizzato per determinare la capacità atletica o il peso corporeo e di conseguenza viene offerta una tariffa peggiore per l'assicurazione malattia, la sua applicazione non è solidale. L'IA può quindi soltanto essere solidale se con l'algoritmo si perseguono obiettivi solidali.

Per finire, quali sono i vantaggi dell'IA per la società?

Jan Kuhlen: Dobbiamo assolutamente lottare contro la connotazione negativa d'intelligenza artificiale. L'automatizzazione dei processi e la riduzione di lavori semplici costituiscono un grande successo, perché si viene a creare ampio spazio per nuovi compiti e creatività. Dobbiamo essere consapevoli del fatto che la rivoluzione digitale è una rivoluzione industriale.

Henrike Schlottmann: Vorrei tornare a parlare del concetto di discriminazione, perché anche in questo campo l'IA ha i suoi lati positivi: gli algoritmi evidenziano una certa discriminazione come viene vissuta nella società. Grazie alla sistematizzazione si può affrontarla attivamente.

Henrike Schlottmann e **Jan Kuhlen** dirigono il gruppo di lavoro che si dedica all'intelligenza artificiale presso «D64 – Zentrum für digitalen Fortschritt» (centro per il progresso digitale), un'associazione non profit di politica digitale. Schlottmann ha un MSci in matematica dello University College London. Per molti anni ha lavorato nella consulenza gestionale internazionale dove ha seguito progetti in ambito di digitalizzazione e innovazione. Oggi, in veste di co-CEO di ProjectTogether sostiene gli imprenditori sociali nell'attuazione dei loro progetti. Kuhlen è laureato in giurisprudenza e sociologia. Con il suo studio legale KUHLEN Partnerschaft von Rechtsanwälten mbB fornisce consulenze a start-up, agenzie e imprese medie in questioni di diritto societario. Jan Kuhlen è membro della commissione del Bundestag dedicata all'intelligenza artificiale e dell'assemblea degli esperti sull'intelligenza artificiale.
