REPORT

AUTOMATING SOCIETY

2020

EDIZIONE ITALIANA

ALGORITHM WATCH

BertelsmannStiftung

IMPRINT

Automating Society Report 2020

Ottobre 2020

Il rapporto è disponibile online all'indirizzo https://automatingsociety.algorithmwatch.org

Edito da
AlgorithmWatch gGmbH
Linienstr. 13
10178 Berlin
Germania

Bertelsmann Stiftung
Carl-Bertelsmann-Str. 256
33311 Gütersloh
Germania

Editor
Fabio Chiusi
Sarah Fischer
Nicolas Kayser-Bril
Matthias Spielkamp

Project manager
Fabio Chiusi

Coordinatore della pubblicazione
Marc Thümmler

Fumettisti
Samuel Daveti
Lorenzo Palloni
Alessio Ravazzani

Layout
Beate Autering
Beate Stangl

Ulteriore revisione dei testi
Leonard Haas

Aggiornato al 30 settembre 2020

Questa pubblicazione è edita su licenza Creative Commons Attribuzione 4.0 Internazionale
https://creativecommons.org/licenses/by/4.0/legalcode.it

Indice

Introduzione 4

Che fare:
raccomandazioni di policy 12

Unione Europea 16

Italia 35

Team 53

Vita quotidiana nella società automatizzata. I sistemi che automatizzano processi decisionali sono diventati mainstream: che fare?

Di Fabio Chiusi

Il materiale contenuto in questo documento è aggiornato al 30 settembre 2020. Non è stato possibile includere avvenimenti e sviluppi occorsi in data posteriore.

In un grigio pomeriggio di agosto, a Londra, gli studenti erano furiosi. Si erano riversati a centinaia in piazza del Parlamento, in protesta, mostrando nei loro cartelli e slogan di schierarsi allo stesso tempo dalla parte di un inusuale alleato, i loro professori, e contro un altrettanto inusuale bersaglio, un algoritmo.

A causa della pandemia di COVID-19, le scuole erano state chiuse già in marzo, in Gran Bretagna. Consci della diffusione del virus in rapida espansione in tutta Europa durante l'estate del 2020, gli studenti ben sapevano che i loro esami di fine anno sarebbero stati cancellati, e la loro valutazione sarebbe di conseguenza — a qualche modo — mutata. Ciò che non avrebbero potuto immaginare, tuttavia, è che migliaia di loro avrebbero finito per ricavarne voti inferiori a quelli attesi.

Gli studenti riuniti in protesta sapevano a chi dare la colpa, come reso evidente dai loro slogan e canti: il sistema di decision-making automatico ("automated decision-making", d'ora in avanti "ADM") adottato dall'Ofqal (Office of Qualifications and Examinations Regulation, l'autorità competente a verificare la correttezza di esami e voti). L'autorità intendeva produrre una migliore valutazione, basata su dati, sia per le valutazioni degli esami per ottenere il GCSE (General Certificate of Secondary Education) che per quelle degli esami "A level" (o General Certificate of Education Advanced Level)[1], in modo che "la distribuzione dei voti segua un andamento simile a quello degli altri anni, così che gli studenti dell'anno in corso non debbano patire uno svantaggio sistemico a seguito delle circostanze presentatesi quest'anno".

Il governo voleva evitare gli eccessi di ottimismo[2] che, secondo le sue valutazioni, si sarebbero prodotti facendo ricorso al solo giudizio umano: in confronto alle serie storiche, i voti sarebbero risultati troppo alti. Ma il tentativo di essere "il più equi possibile, equi per gli studenti che non hanno potuto dare gli esami quest'estate" aveva invece fallito miseramente, e in quel grigio giorno di protesta d'agosto, gli studenti continuavano ad accorrere, cantare, mostrare cartelli che esprimevano un urgente bisogno di giustizia sociale. Alcuni di loro erano semplicemente disperati, altri si erano lasciati andare al pianto.

"Basta rubarci il futuro", recitava un cartello, facendo il verso alle proteste per la giustizia climatica di Fridays for the Future. Altri, tuttavia, erano più specificamente tarati sulle falle del sistema di ADM per l'assegnazione dei voti: "Valuta il mio lavoro, non il mio codice postale", siamo "studenti, non statistiche", recavano scritto, denunciando così i risultati discriminatori del sistema[3].

Un canto levatosi in seguito ha finito, tuttavia, per definire il futuro stesso della protesta: intonava "*Fuck the algorithm*", "fanculo l'algoritmo". Terrorizzati che il governo stesse automatizzando con noncuranza — e in modo opaco — il loro futuro, indipendentemente dalle loro reali abilità e sforzi, gli studenti avevano preso a gridare il loro diritto di non vedersi indebitamente restringere le opportunità di vita da un insieme di righe di codice mal programmate. Reclamavano diritto di parola, e ciò che avevano da dire, in effetti, andava ascoltato.

Perché gli algoritmi non sono né "neutrali" né oggettivi, anche se tendiamo a pensarlo. Replicano, invece, gli assunti e le credenze di chi decide di adottarli e programmarli. È sempre un umano dunque, non "gli algoritmi" o i sistemi di ADM, a essere responsabile sia delle buone che della cattive decisioni algoritmiche — o almeno, così dovrebbe essere. La macchina sarà pure sinistra, ma il fantasma al suo interno è sempre umano. E gli esseri umani sono complicati, perfino più degli algoritmi.

Gli studenti in protesta non erano ingenui al punto di credere che tutte le loro preoccupazioni dipendessero esclusivamente dall'algoritmo, in ogni caso. Non stavano intonando cori contro "l'algoritmo" in un raptus di determinismo tecnologico: erano piuttosto motivati dall'urgenza di promuovere e proteggere la giustizia sociale. Da questo punto di vista, le loro proteste ricordano piuttosto quelle dei Luddisti. Proprio come il movimento per i diritti dei lavoratori che, nel XIX secolo, distruggeva telai meccanici e altre macchine da lavoro industriale, gli studenti sanno che i sistemi di ADM sono questione di potere, e non dovrebbero venire scambiati per tecnologie "oggettive". Per questo cantavano "giustizia alla classe lavoratrice", chiedendo le dimissioni del ministro della Salute e ritraendo il sistema di ADM incriminato come "classismo in purezza" e "classismo spudorato".

1 Nel sistema scolastico del Regno Unito, il "GCSE" è un esame affrontato dagli studenti di circa 16 anni; gli esami "A level" si tengono di norma dopo un ulteriore biennio di studio, e servono a decidere l'ingresso nel sistema educativo universitario.

2 "La letteratura suggerisce che, nello stimare la probabilità dei voti che gli studenti otterranno, gli insegnanti tendono a essere ottimisti (anche se non in tutti i casi)", scrive l'Ofqual, cfr. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/909035/6656-2_-_Executive_summary.pdf

3 Cfr. il capitolo sulla Gran Bretagna per i dettagli.

Infine, gli studenti sono riusciti ad abolire il sistema che avrebbe messo le loro opportunità professionali e di vita in pericolo: in una incredibile giravolta, il governo britannico ha infatti dismesso il sistema di ADM rivelatosi fallace, e deciso di affidarsi al giudizio degli insegnanti.

Ma c'è altro in questa storia, oltre alla vittoria dei manifestanti. L'esempio sottolinea infatti come sistemi mal progettati, implementati e controllati che riproducono bias e discriminazioni umane non possano mettere a frutto i potenziali benefici dei sistemi di ADM, per esempio in termini di comparabilità ed equità.

Più chiaramente rispetto a molte battaglie del passato, questa protesta rivela che non stiamo più solamente "automatizzando la società", ma che l'abbiamo già automatizzata.

Finalmente, qualcuno se ne è accorto.

/ Da "Automating Society" alla società automatizzata

Quando, nel gennaio 2019, abbiamo lanciato la prima edizione di questo rapporto, abbiamo deciso di chiamarlo "Automating Society", cioè "automatizzando la società", perché allora i sistemi di ADM reperibili in Europa erano principalmente nuovi, sperimentali e sconosciuti. Soprattutto, erano l'eccezione, piuttosto che la norma.

La situazione è cambiata rapidamente. Come mostrato chiaramente nei molteplici esempi raccolti per questo rapporto dal nostro straordinario network di ricercatori, l'adozione di sistemi di ADM è fortemente incrementata in poco più di un anno. I sistemi di ADM riguardano oggi ogni tipo di attività umana e, in particolare, la distribuzione di servizi essenziali a milioni di cittadini europei — nonché il loro effettivo accesso ai propri diritti.

L'opacità testarda che continua a circondare l'uso sempre crescente di sistemi di ADM ha reso anche più urgente moltiplicare, di conseguenza, i nostri sforzi. Per questo abbiamo aggiunto quattro paesi (Estonia, Grecia, Portogallo e Svizzera) ai dodici già analizzati nell'edizione precedente di questo rapporto, portando il totale a sedici paesi. Anche se ben lontana da rappresentare un'analisi esaustiva, ciò ci consente di fornire un ritratto più ampio dello scenario dell'ADM in Europa. Considerando l'impatto che questi sistemi possono avere sulla vita di ogni giorno, e quanto in profondità sfidino le nostre intuizioni — quando non le norme e regole — circa il rapporto tra democrazia e automazione, crediamo si tratti di uno sforzo indispensabile.

Ciò è vero a maggior ragione nel contesto della pandemia di COVID-19, un momento storico in cui abbiamo testimoniato l'adozione (perlopiù affrettata) di una pletora di sistemi di ADM mirati a contribuire alla salute pubblica attraverso dati e automazione. Abbiamo ritenuto il fenomeno talmente importante da farne l'oggetto di un "rapporto-anteprima" (*preview report*), pubblicato[4] ad agosto 2020 sempre all'interno del progetto 'Automating Society'.

Perfino in Europa, quando si tratta di adottare sistemi di ADM, il limite è la fantasia. Si pensi anche solo ad alcuni dei casi raccontati in questo rapporto, che vanno ad aggiungersi ai tanti — dal welfare all'educazione, dal sistema sanitario a quello giudiziario — di cui abbiamo già dato conto nell'edizione precedente. Nelle pagine che seguono, e per la prima volta, forniremo aggiornamenti e sviluppi su questi casi in tre modi: tramite articoli giornalistici, attraverso una sezione di ricerca che elenca i diversi esempi raccolti, e infine con storie illustrate. Abbiamo ritenuto che questi sistemi di ADM siano — e saranno sempre più — talmente importanti nelle vite di ciascuno di noi da richiedere ogni possibile sforzo per comunicare come funzionino, e *cosa ci facciano* davvero, così da raggiungere ogni possibile tipologia di pubblico. Dopotutto, i sistemi di ADM ci riguardano tutti.

O almeno, dovrebbero. Abbiamo visto, per esempio, come un nuovo servizio automatizzato, e proattivo, distribuisca i bonus famiglia in Estonia. I genitori non devono nemmeno più fare domanda: lo Stato registra semplicemente tutte le informazioni di un neonato fin dalla nascita, raccogliendole in diversi database. Il risultato è che i genitori ricevono il bonus automaticamente, se ne hanno diritto.

In Finlandia, l'identificazione di fattori di rischio individuali correlati all'esclusione sociale dei giovani è automatizzata attraverso uno strumento sviluppato dal gigante giapponese Fujitsu. In Francia, i dati dei social network possono essere analizzati per addestrare algoritmi di machine learning che vengono impiegati per scovare frodi fiscali.

L'Italia sta sperimentando la "giustizia predittiva", che fa ricorso all'automazione per aiutare i giudici a individuare, nella giurisprudenza passata, tendenze utili a valutare

4 'Automated Decision-Making Systems in the COVID-19 Pandemic: A European Perspective', https://algorithmwatch.org/en/project/automating-society-2020-covid19/

il caso in esame. E, in Danimarca, il governo ha provato a controllare ogni click di tastiera e mouse sui computer degli studenti durante gli esami di fine anno, causando — anche qui — una considerevole protesta studentesca che ha portato all'abbandono del sistema, per il momento.

/ È tempo di correggere i torti dell'ADM

In linea di principio, i sistemi di ADM hanno il potenziale di arrecare beneficio alle vite dei cittadini— processando per esempio enormi moli di dati, accompagnando le persone nei loro processi decisionali, e fornendo loro applicazioni personalizzate.

In pratica, tuttavia, abbiamo reperito solo un numero esiguo di casi che dimostri in modo convincente tale impatto positivo.

Per esempio, il sistema VioGén, adottato in Spagna sin dal 2007 per valutare il rischio di violenze domestiche, pur se lontano dalla perfezione mostra "indici di prestazione ragionevoli" e ha contribuito a proteggere diverse donne da abusi.

In Portogallo, un sistema centralizzato e automatico adottato per combattere frodi associate a prescrizioni mediche ha apparentemente ridotto le frodi dell'80% in un solo anno. Un sistema simile, usato in Slovenia contro le frodi fiscali, si è dimostrato utile agli ispettori, a quanto dichiara il fisco sloveno[5].

Il riconoscimento facciale, quasi assente dall'edizione 2019, sta venendo sperimentato e adottato a ritmi allarmanti in tutta Europa.

Se si amplia lo sguardo allo stato attuale dei sistemi di ADM in Europa, si scopre che gli esempi positivi, che arrecano chiari benefici, sono rari. Nel corso di tutto il rapporto, si descrive piuttosto come la stragrande maggioranza degli usi tenda ad aumentare i rischi per i cittadini, invece di essere loro d'aiuto. Ma, per giudicare davvero i reali impatti positivi e negativi di questi sitemi, abbiamo bisogno di maggiore trasparenza circa i loro scopi, e di più dati sul funzionamento dei sistemi di ADM sperimentati e adottati.

Il messaggio per i decisori politici non potrebbe essere più chiaro. Se vogliamo davvero sfruttare al massimo il potenziale dei sistemi di ADM, rispettando insieme i diritti umani e la democrazia, il tempo di intervenire, renderli trasparenti e correggerne i torti è ora.

/ Riconoscimento facciale, riconoscimento facciale ovunque

Strumenti diversi stanno venendo adottati in paesi diversi. Una tecnologia, tuttavia, è oramai comune a molti: il riconoscimento facciale. È ciò che può essere definito lo sviluppo più nuovo, rapido e preoccupante tra quelli evidenziati in questo rapporto. Il riconoscimento facciale, quasi assente dall'edizione 2019, sta venendo sperimentato e adottato a ritmi allarmanti in tutta Europa. Nell'anno e poco più trascorso dal nostro ultimo rapporto, il riconoscimento facciale ha fatto il suo ingresso in scuole, stadi, aeroporti, e perfino nei casinò. È stato così usato come strumento di polizia predittiva, per arrestare criminali, contro il razzismo e, in risposta alla pandemia di COVID-19, per garantire il rispetto delle norme di distanziamento sociale, sia via app che attraverso videosorveglianza "intelligente".

L'adozione di nuovi sistemi di riconoscimento facciale prosegue nonostante il cumularsi delle prove scientifiche circa la loro scarsa accuratezza. E ogniqualvolta emerga un problema, i loro proponenti cercano di trovare un modo di aggirarlo. In Belgio, un sistema di riconoscimento facciale utilizzato dalla polizia è ancora "parzialmente attivo", anche se un divieto temporaneo è stato emanato da un organo di controllo, l'Oversight Board for Police Information. E, in Slovenia, l'uso di tecnologie di riconoscimento facciale da parte della polizia è stato legalizzato cinque anni dopo che gli agenti avevano cominciato a utilizzarle.

5 Cfr. il capitolo sulla Slovenia per i dettagli.

Questa tendenza, se incontrastata, rischia di normalizzare l'idea di essere costantemente osservati senza avere alcuna trasparenza su chi osserva, finendo così per cristallizzare uno status quo di sorveglianza di massa pervasiva. Ecco perché, su questo, molti all'interno della comunità delle organizzazioni per i diritti civili avrebbero accolto con favore una risposta politica molto più aggressiva da parte delle istituzioni europee[6].

Perfino il proprio sorriso è ora parte di un sistema di ADM sperimentato in alcune banche polacche: più l'impiegato sorride, migliore è la ricompensa. E a essere sotto controllo non sono solo i volti. In Italia, un sistema di sorveglianza sonora è stato proposto come strumento anti-razzismo in tutti gli stadi di calcio.

/ Le scatole nere sono ancora scatole nere

Un risultato allarmante del nostro rapporto è che, mentre le cose sono rapidamente cambiate per quanto riguarda il tasso di adozione dei sistemi di ADM, lo stesso non si può dire della loro trasparenza. Nel 2015, il docente della Brooklyn Law School, Frank Pasquale, popolarizzò l'idea che una società connessa basata su sistemi algoritmici opachi fosse una "black box society", ossia una società fatta di "scatole nere", i cui contenuti e dinamiche di funzionamento sono inconoscibili. Cinque anni dopo, sfortunatamente, la metafora ancora regge — e si applica a tutti i paesi studiati in questo rapporto, senza eccezioni: non c'è abbastanza trasparenza sui sistemi di ADM, né nel settore pubblico, né in quello privato. La Polonia addirittura obbliga all'opacità, con la legge che ha introdotto il suo sistema automatizzato per riconoscere conti bancari usati per attività illegali ("STIR"). La legge stabilisce che rivelare gli algoritmi e gli indicatori di rischio adottati può comportare fino a cinque anni di carcere.

Se da un lato rigettiamo l'idea che tutti questi sistemi siano intrinsecamente malvagi — abbracciamo, al contrario, una prospettiva pragmatica, basata su fatti ed evidenze scientifiche (*evidence-based*) — è dall'altro indubbiamente male essere incapaci di valutarne il funzionamento e l'impatto sulla base di conoscenze accurate e fattuali. Anche solo perché l'opacità ostacola gravemente la capacità di raccogliere prove necessarie a formulare un giudizio informato sull'adozione stessa dei sistemi di ADM.

Quando vi si aggiungono le difficoltà che entrambi i nostri ricercatori e giornalisti hanno incontrato nell'accedere a qualunque dato realmente significativo su questi sistemi, si comprende quanto lo scenario dipinto sia problematico per chiunque desideri tenerli sotto controllo, e garantire che lo sviluppo dei sistemi di ADM sia compatibile con i diritti fondamentali, lo stato di diritto, e la democrazia.

/ Sfidare lo status quo algoritmico

Di fronte a tutto questo, come sta reagendo l'Unione Europea? Anche se i documenti strategici prodotti dalla Commissione UE, sotto la guida di Ursula Von der Leyen, fanno più generico riferimento all'"intelligenza artificiale" piuttosto che parlare direttamente di sistemi di ADM, ciò non significa che non contengano lodevoli intenzioni, a partire da quella di promuovere e realizzare una "AI degna di fiducia" ("trustworthy AI") che veda al centro la persona umana ("people first")[7].

Eppure di fatto, come descritto nel capitolo sull'UE, l'approccio complessivo dell'Unione dà la priorità all'imperativo, commerciale e geopolitico, di guidare la "rivoluzione dell'AI" piuttosto che a quello di assicurarsi che i suoi prodotti siano coerenti con le tutele democratiche, una volta adottati come strumenti di policy.

Questa mancanza di coraggio politico, specialmente evidente nella decisione di accantonare l'idea di una moratoria sull'uso di tecnologie di riconoscimento facciale dal vivo in luoghi pubblici nel suo pacchetto di norme sull'AI, è sorprendente, in particolar modo in un frangente storico in cui diversi stati membri stanno testimoniando un numero crescente di cause — e sconfitte — giudiziarie per sistemi di ADM da loro adottati troppo frettolosamente, finendo per impattare in modo negativo sui diritti dei cittadini.

Storico è un caso proveniente dall'Olanda, dove degli attivisti per i diritti civili sono riusciti a portare in giudizio un sistema automatizzato opaco e invasivo, chiamato SyRI, che avrebbe dovuto individuare casi di frode al sistema del welfare, e vincere. Non solo, infatti, a febbraio 2020 la corte dell'Aia ha sospeso il sistema, ritenendolo in violazione della Convenzione europea dei diritti dell'uomo. Il caso ha anche istituito un precedente: secondo la sentenza, i governi hanno la "speciale responsabilità" di salvaguardare i diritti umani, qualora adottino tali sistemi di ADM. Fornire

6 Come argomentato in dettaglio nel capitolo sull'UE.

7 Si veda il capitolo sull'UE, e in particolare la sezione dedicata al 'Libro bianco sull'AI' della Commissione Europea.

Un risultato allarmante del nostro rapporto è che, mentre le cose sono rapidamente cambiate per quanto riguarda il tasso di adozione dei sistemi di ADM, lo stesso non si può dire della loro trasparenza.

la tanto agognata trasparenza algoritmica ne è considerata parte cruciale.

Più in generale, dall'uscita del nostro primo rapporto, i media e gli attivisti della società civile si sono imposti quali forze trainanti verso la responsabilizzazione dei sistemi di ADM. In Svezia, per esempio, sono stati dei giornalisti a forzare la pubblicazione del codice del "sistema Trelleborg", creato per prendere decisioni completamente automatizzate in relazione a domande di sussidi sociali. A Berlino, il progetto-pilota per un sistema di riconoscimento facciale nella stazione ferroviaria di Südkreuz non ha condotto a una vera e propria implementazione in tutto il paese solo grazie alla rumorosa opposizione degli attivisti — talmente rumorosa da influenzare le posizioni dei principali partiti e, da ultimo, l'agenda politica del governo.

Gli attivisti greci di Homo Digitalis hanno dimostrato che nessun reale viaggiatore ha mai partecipato alle sperimentazioni del sistema chiamato 'iBorderCtrl', un progetto finanziato dall'UE il cui obiettivo era utilizzare sistemi di ADM per i controlli ai confini, chiarendo così che le effettive capacità di molti tra questi sistemi sono di frequente grandemente esagerate. Nel frattempo, in Danimarca, un sistema di profilazione per l'identificazione precoce di rischi associati a famiglie e figli con "fragilità" (il cosiddetto "modello Gladsaxe") veniva a sua volta fermato grazie al lavoro di accademici, giornalisti e dell'Autorità nazionale per la protezione dei dati personali (DPA).

Le Authority per la privacy hanno giocato un ruolo importante anche in altri paesi. In Francia, la DPA nazionale ha stabilito che due progetti per la sorveglianza sonora e il riconoscimento facciale negli istituti scolastici superiori fossero entrambi illegali. In Portogallo, ha rifiutato di approvare l'adozione di un sistema di videosorveglianza nelle città di Leiria e Portimão, perché ritenuto in violazione del principio di proporzionalità, e perché avrebbe costituito l'equivalente di un "monitoraggio sistematico su larga scala", un "tracciamento delle persone e delle loro abitudini e comportamenti", e di una "identificazione degli individui a partire da dati correlati a caratteristiche fisiche". In Olanda, poi, l'autorità Garante dei dati personali ha chiesto più trasparenza circa gli algoritmi predittivi utilizzati dalle agenzie governative.

Da ultimo, alcuni paesi hanno fatto ricorso alla figura del difensore civico (*ombudsperson*), a caccia di consigli. In Danimarca, il suo supporto ha contribuito a sviluppare strategie e linee guida per un uso etico dei sistemi di ADM nel settore pubblico. In Finlandia, il vice-difensore civico parlamentare ha considerato illegale la valutazione automatica del livello di tassazione.

E ciononostante, data la continua diffusione di tali sistemi in tutta Europa, viene da chiedersi: questo livello di controllo è sufficiente? Quando per esempio il difensore civico polacco ha messo in discussione la legalità del sistema di riconoscimento dei sorrisi adottato in una banca (e menzionato sopra), la decisione non ha impedito una ulteriore e suc-

cessiva sperimentazione nella città di Sopot, né raffreddato l'interesse di diverse aziende, ancora decise ad adottarlo.

/ Mancano auditing, conseguenze, capacità e spiegazioni

L'attivismo è principalmente una misura reattiva. Il più delle volte, gli attivisti possono reagire solo quando un sistema di ADM sta già venendo sperimentato, o è addirittura già stato adottato. Nel tempo necessario a organizzare una risposta, i diritti dei cittadini rischiano di essere già stati indebitamente intaccati. Ciò può accadere perfino in presenza delle forme di protezione garantite, in molti casi, dalla normativa europea e degli stati membri. È per questa ragione che garanzie proattive — e preventive, anteriori alla loro sperimentazione e adozione — sono così importanti per l'effettiva salvaguardia dei diritti dei cittadini.

Eppure, perfino nei paesi in cui una qualche forma di legislazione che includa tutele proattive è in vigore, l'effettiva applicazione della legge, molto semplicemente, non si verifica. In Spagna, per esempio, l'"azione amministrativa automatizzata" è codificata per legge, con annessi obblighi in termini di controllo qualità e supervisione, così come di auditing del sistema informatico e del suo codice sorgente. La Spagna si è anche dotata di una legge per l'accesso all'informazione (*freedom of information*). Eppure, nonostante queste leggi, solo raramente, scrive il nostro ricercatore, gli organi pubblici condividono informazioni dettagliate circa i sistemi di ADM che utilizzano. Allo stesso modo, in Francia esiste una legge che dal 2016 obbliga alla trasparenza algoritmica ma, di nuovo, senza esito.

Nemmeno portare un algoritmo di fronte a una corte giudiziaria, sulla base di specifiche disposizioni contenute in una legge per la trasparenza algoritmica, è abbastanza per proteggere davvero i diritti degli utenti. Come dimostrato dal caso francese dell'algoritmo di Parcoursup per la selezione di studenti universitari[8], le eccezioni per mettere al riparo una amministrazione pubblica da ogni forma di *accountability* si possono ricavare a piacimento.

Ciò è particolarmente problematico quando vi si aggiunga un contesto caratterizzato dalla endemica mancanza, nella pubblica amministrazione, di capacità e competenze riguardanti i sistemi di ADM da tempo lamentata da molti dei nostri ricercatori. E come potrebbero i pubblici ufficiali spiegare o fornire alcuna reale trasparenza su sistemi che non comprendono?

Di recente, alcuni paesi hanno tentato di affrontare il problema. L'Estonia, per esempio, ha predisposto un centro per competenze (*competence center*) rilevanti per i sistemi di ADM, per meglio comprendere come potrebbero venire usati per sviluppare nuovi servizi pubblici e, più nello specifico, informare le operazioni del Ministero per gli Affari economici e le comunicazioni, e della Cancelleria di Stato per lo sviluppo dell'e-government. Anche la Svizzera ha proposto la creazione di una "rete di competenze" (*competence network*) all'interno della più ampia cornice della strategia nazionale, chiamata "Svizzera Digitale".

Ma ciononostante, la ben nota mancanza di alfabetismo digitale resta un problema per buona parte della popolazione in diversi paesi europei. Per di più, è difficile chiedere il rispetto di diritti che non si sa di avere. Le proteste in Gran Bretagna e altrove, insieme a diversi scandali pubblici a base di sistemi di ADM[9], hanno di certo contribuito a innalzare il livello di consapevolezza sia dei rischi che delle opportunità derivanti dall'automatizzazione della società. Ma per quanto in crescita, questa consapevolezza muove ancora solo i primi passi in molti paesi.

I risultati della nostra ricerca sono chiari: per quanto i sistemi di ADM già influenzino ogni sorta di attività e giudizio, è ancora loro concesso di essere principalmente adottati senza alcun reale dibattito democratico. Inoltre è la norma, piuttosto che l'eccezione, osservare meccanismi di tutela e controllo in palese ritardo rispetto all'adozione dei sistemi che dovrebbero controllare — se e quando esistono.

Nemmeno lo scopo di questi sistemi viene comunemente giustificato o spiegato alle popolazioni che ne sono affette, né tantomeno vengono illustrati i benefici che ne dovrebbero loro derivare. Si pensi al servizio proattivo "AuroraAI" in Finlandia: dovrebbe identificare automaticamente gli "eventi di una vita" (*life events*), come riportano i nostri ricercatori finlandesi. Nella mente dei proponenti dovrebbe funzionare come una sorta di "tata", capace di aiutare i cittadini a soddisfare i loro bisogni di servizi pubblici in concomitanza con precisi eventi della propria vita — per esempio, cambiare residenza, cambiare stato nelle relazioni familiari, etc. Una forma di spiacevole "spinta gentile" (*nudging*) potrebbe essere all'opera in questo caso, scrivono i ricercatori, stan-

8 Cfr. il capitolo sulla Francia.

9 Si pensi al fiasco dell'algoritmo de "La buona scuola" in Italia, cfr. il relativo capitolo.

do a dire che, invece di avvantaggiare le persone, il sistema potrebbe finire per fare l'esatto opposto, suggerendo alcune decisioni o limitando le opzioni a disposizione di un individuo già con il proprio design o la sua stessa architettura.

È di conseguenza anche più importante sapere cosa, più di preciso, stia venendo "ottimizzato" in termini di servizi pubblici: "è massimizzato l'uso del servizio, sono minimizzati i costi, o viene migliorato il benessere del cittadino?", chiedono i ricercatori. "Su quale insieme di criteri si basano queste decisioni, e chi li stabilisce?" Il semplice fatto che non si abbiano risposte a queste fondamentali domande la dice lunga sul reale grado di partecipazione e trasparenza concesso, perfino per un sistema di ADM tanto invasivo.

/ La trappola tecno-soluzionista

C'è una più ampia giustificazione ideologica per tutto questo. È chiamata "soluzionismo tecnologico", e affligge ancora in profondità il modo in cui i sistemi di ADM vengono studiati e sviluppati. Anche se il termine è da tempo criticato come sinonimo di una ideologia fallace, che concepisce ogni problema sociale come un "bug" in attesa di essere aggiustato ("fix") tramite la tecnologia[10], questa retorica è ancora ampiamente sfruttata — sia nei media che in ambienti di policy — per giustificare l'adozione acritica di tecnologie di automazione nella vita pubblica.

Quando dipinti come "soluzioni", i sistemi di ADM fanno immediatamente ingresso nei territori meglio descritti dalla Terza Legge di Arthur C. Clarke: quelli in cui sono indistinguibili dalla magia. Ed è difficile, se non impossibile, imporre delle regole alla magia, così come renderla trasparente o spiegarla. Ciò che si vede è la mano frugare nel cappello e, di conseguenza, l'apparire di un coniglio, ma il procedimento è e *deve rimanere* una "black box", inconoscibile.

Diversi tra i ricercatori coinvolti nel progetto 'Automating Society' hanno individuato in questa ideologia l'errore fondamentale che informa la logica con cui vengono concepiti i sistemi di ADM descritti nel rapporto. Ciò implica anche, come mostrato nel capitolo sulla Germania, che gran parte delle critiche a tali sistemi vengano descritte come puro e semplice rigetto dell'"innovazione" tutta, mentre i sostenitori dei diritti digitali non sarebbero che "neo-luddisti". Ciò tuttavia non solo dimentica la realtà storica del movimento luddista, che si occupava di politiche del lavoro e non di mere tecnologie, ma minaccia inoltre, in un senso perfino più elementare, di compromettere l'efficacia dei meccanismi di controllo ipotizzati.

In un frangente storico in cui l'industria dell'AI sta testimoniando l'emergenza di un settore lobbistico particolarmente "vitale", a partire dalla Gran Bretagna, ciò potrebbe tradursi nell'adozione di comodo di mere linee guida per l'etica ("ethics-washing", il tentativo di prevenire norme e regole con un surplus di autoregolamentazione volontaria) e in altre risposte di policy inefficaci e strutturalmente inadeguate ad affrontare le implicazioni sui diritti umani dei sistemi di ADM. Ciò significherebbe, in ultima analisi, sposare l'assunto secondo cui siamo noi esseri umani a doverci adattare ai sistemi di ADM, molto più che questi ultimi a dover essere modellati secondo i principi delle società democratiche.

Per contrastare questa narrativa, non dovremmo trattenerci dal porre domande fondamentali: per esempio, se i sistemi di ADM siano compatibili con la democrazia e possano essere impiegati per arrecare beneficio alla società tutta, e non solo ad alcuni. Potrebbe darsi, infatti, che certe attività umane — per dirne una tipologia, quelle collegate al welfare — non debbano essere assoggettate ad automazione, o che certe tecnologie — a partire dal riconoscimento facciale dal vivo in luoghi pubblici — non debbano essere incentivate nella ricerca, senza fine, della "leadership sull'AI", ma al contrario messe al bando.

Anche più importante è rigettare qualunque cornice ideologica ci impedisca di porre simili domande. Al contrario: ciò di cui abbiamo bisogno ora è che cambino alcune scelte politiche concrete — così che sia consentito un più approfondito scrutinio di questi sistemi. Nella sezione successiva elencheremo le principali richieste emerse dai nostri risultati di ricerca. Speriamo che siano ampiamente dibattute, e infine implementate.

Solo attraverso un dibattito informato, inclusivo, basato su evidenze scientifiche — un dibattito davvero democratico, dunque — riusciremo a trovare il giusto bilanciamento tra i benefici che i sistemi di ADM posso arrecare — e arrecano — in termini di rapidità, efficienza, equità, migliore prevenzione e accesso ai servizi pubblici, e le sfide che pongono ai diritti di tutti.

10 Si veda in proposito Evgeny Morozov (2014), *To Save Everything, Click Here. The Folly of Technological Solutionism*, Public Affairs, https://www.publicaffairsbooks.com/titles/evgeny-morozov/to-save-everything-click-here/9781610393706/

Che fare: raccomandazioni di policy

Alla luce dei risultati dettagliati nell'edizione 2020 del rapporto 'Automating Society', consigliamo ai policy-maker nel Parlamento Europeo e nei Parlamenti degli stati membri, alla Commissione UE, ai governi nazionali, ai ricercatori, alle organizzazioni della società civile (organizzazioni di *advocacy*, fondazioni, sindacati, etc.), e al settore privato (aziende e associazioni commerciali) di adottare il seguente insieme di interventi di policy. Queste raccomandazioni hanno l'obiettivo di meglio assicurare che i sistemi di ADM attualmente adottati e in via di adozione in tutta Europa siano effettivamente coerenti con il rispetto dei diritti umani e delle regole democratiche:

1. Più trasparenza per i sistemi di ADM

Senza la possibilità di sapere precisamente come, perché e a quali fini i sistemi di ADM vengano adottati, tutti gli altri sforzi per riconciliarli con i diritti fondamentali sono destinati a fallire.

/ Creare registri pubblici per i sistemi di ADM utilizzati nel settore pubblico

Chiediamo, di conseguenza, che venga adottata una legge di rango europeo che obblighi gli stati membri alla creazione di registri pubblici per i sistemi di ADM usati nel settore pubblico.

Dovrebbero inoltre essere accompagnati da obblighi di legge, per i responsabili dei sistemi di ADM, riguardanti la trasparenza e la documentazione dello scopo del sistema, una spiegazione del modello adottato (logica inclusa), e informazioni su chi l'ha sviluppato. Tutte queste informazioni devono essere rese disponibili in formati facilmente leggibili e accessibili, inclusi dati digitali strutturati e basati su protocolli standardizzati.

Le autorità pubbliche dovrebbero avere la responsabilità specifica di rendere trasparenti le componenti operative dei sistemi di ADM adottati nelle amministrazioni pubbliche — come sottolineato da un recente reclamo amministrativo in Spagna, che argomenta che "ogni sistema di ADM usato da un'amministrazione pubblica dovrebbe essere reso pubblico di default". Se confermato, il giudizio potrebbe diventare un precedente, in Europa.

/ Introdurre schemi legalmente vincolanti per l'accesso ai dati, a supporto della ricerca nell'interesse pubblico

Aumentare il grado di trasparenza fornito non richiede solamente la divulgazione di informazioni circa scopo, logica e creatore di un sistema, né la capacità di analizzarne nel dettaglio, e mettere alla prova, gli input e gli output. Richiede infatti anche che i dati attraverso cui i suoi algoritmi vengono addestrati, e i risultati da loro prodotti, siano resi accessibili a ricercatori indipendenti, giornalisti e organizzazioni della società civile, nell'interesse pubblico.

Ecco perché suggeriamo l'introduzione di schemi solidi e legalmente vincolanti per l'accesso ai dati, esplicitamente mirati a supportare e promuovere la ricerca a fini di pubblico interesse, e nel pieno rispetto della normativa sulla protezione dei dati e la privacy.

Facendo tesoro delle migliori esperienze (*best practices*) a livello nazionale ed europeo, questi schemi a più livelli dovrebbero includere sistemi di sanzioni, di tutele (*checks and balances*), e revisioni periodiche. Come illustrato dalle partnership per la condivisione dei dati con soggetti privati, ci sono legittime preoccupazioni in termini di privacy degli utenti e di possibile deanonimizzazione di certe tipologie di dati.

I policy-maker dovrebbero fare propri gli insegnamenti derivanti dai modelli di condivisione dei dati sanitari, così da rendere più semplice dare accesso privilegiato a certe tipologie di dati, più granulari, e al contempo garantire che i dati personali siano protetti in modo adeguato (per esempio, attraverso ambienti operativi sicuri).

E se ottenere uno schema di responsabilizzazione efficace richiede di certo un accesso trasparente ai dati in possesso delle piattaforme, quest'ultimo è un requisito indispensabile anche all'efficacia di svariati approcci di auditing.

2. Creare uno schema per una reale responsabilizzazione in tema di sistemi di ADM

Come dimostrato da quanto documentato per Spagna e Francia, perfino quando la trasparenza di un sistema di

ADM diventa norma di legge e/o delle informazioni al suo riguardo vengono effettivamente divulgate, ciò non necessariamente comporta reale *accountability.* Affinché disposizioni di legge e altri requisiti vengano effettivamente rispettati, servono passi ulteriori.

/ Sviluppare e adottare approcci per un effettivo audit dei sistemi algoritmici

Per garantire che la trasparenza sia reale, c'è bisogno di completare il primo passo — la creazione di un registro pubblico — con processi che consentano un effettivo *audit* dei sistemi algoritmici.

Il termine "auditing", in italiano associato a "revisione", "verifica" e "controllo", è di largo utilizzo, ma non c'è consenso intorno a una sua comune definizione. Per noi, in questo contesto va compreso in accordo con la definizione dell'ISO: un "processo sistematico, indipendente e documentato per ottenere evidenze oggettive e valutarle oggettivamente, così da determinare in che misura i criteri dell'audit sono stati rispettati"[11].

Non disponiamo ancora di risposte soddisfacenti per tutte le complesse domande[12] sollevate dall'auditing dei sistemi algoritmici. Tuttavia, i nostri risultati mostrano chiaramente il bisogno di trovarle attraverso un ampio processo di coinvolgimento dei portatori di interesse, e tramite un lavoro di ricerca approfondito e specifico.

Dovrebbero essere sviluppati sia criteri che processi appropriati a costruire un sistema di auditing efficace, attraverso un approccio multi-stakeholder che tenga attivamente in considerazione l'effetto sproporzionato che i sistemi di ADM hanno sui gruppi sociali più vulnerabili — e ne solleciti, di conseguenza, la partecipazione.

Chiediamo ai policy-maker, di conseguenza, di dare inizio a questi processi con i portatori di interesse, così da chiarire le domande illustrate, e rendere disponibili le fonti di finanziamento necessarie a consentire la partecipazione degli stakeholder finora non rappresentati in modo adeguato.

Chiediamo inoltre la predisposizione di risorse adeguate a supportare e finanziare progetti di ricerca sullo sviluppo di modelli efficaci di auditing per i sistemi algoritmici.

/ Promuovere le organizzazioni della società civile a watchdog dei sistemi di ADM

I risultati della nostra ricerca indicano chiaramente che il lavoro delle organizzazioni della società civile è cruciale per sfidare efficacemente l'opacità dei sistemi di ADM. Attraverso ricerca e attivismo, e spesso in cooperazione con istituzioni accademiche e giornalisti, tali entità sono negli ultimi anni ripetutamente intervenute nei dibattiti politici su automazione e democrazia, riuscendo in molti casi ad assicurare che l'interesse pubblico e i diritti fondamentali fossero debitamente tenuti in considerazione, sia prima che dopo l'adozione di sistemi di ADM, in diversi paesi europei.

La società civile dovrebbe essere dunque supportata, in qualità di "cane da guardia" (*watchdog*) della società automatizzata. Come tale, la società civile è parte integrante di qualunque schema di *accountability* dei sistemi di ADM voglia dirsi efficace.

/ Mettere al bando il riconoscimento facciale capace di sorveglianza di massa

Non tutti i sistemi di ADM sono ugualmente pericolosi, e un approccio regolatorio basato sul rischio, come quello tedesco o della UE, lo riconosce correttamente. Ma se l'intento è fornire reale *accountability* per sistemi identificati come rischiosi, vanno insieme creati meccanismi efficaci di controllo e implementazione. Ciò è anche più importante per sistemi considerati "ad alto rischio" di violare i diritti degli utenti.

11 https://www.iso.org/obp/ui/#iso:std:iso:19011:ed-3:v1:en

12 Pensando a potenziali modelli di auditing algoritmico, emergono diverse domande. 1) Chi o cosa (servizi, piattaforme, prodotti) dovrebbe essere oggetto dell'auditing? Come personalizzare i sistemi di auditing alla tipologia di piattaforma o servizio? 2) Quando dovrebbe essere responsabilità di una istituzione pubblica (a livello europeo, nazionale o locale), e quando invece può essere fatto da soggetti privati e da esperti (in campo commerciale, nella società civile o nella ricerca)? 3) Come chiarire la distinzione tra valutare un impatto ex-ante (i.e. nella fase di design) ed ex-post (i.e. quando è già in funzione), e le rispettive sfide? 4) Come valutare i trade-off tra virtù e vizi dell'auditing (per esempio, semplicità, generalità, applicabilità, precisione, flessibilità, interpretabilità, privacy, efficacia di una procedura di auditing possono essere in conflitto)?; 5) Quali informazioni devono essere disponibile affinché un audit sia efficace e affidabile (codice sorgente, dati di training, documentazione)? Gli ispettori necessitano di avere accesso fisico ai sistemi durante il loro funzionamento per compiere un audit efficace? 6) Quali obblighi di produrre prove sono necessari e proporzionati per venditori e fornitori di servizi? 7) Come possiamo assicurare che l'auditing sia possibile? C'è bisogno che i requisiti dell'auditing vengano considerati già nella fase di modellazione (*design*) del sistema algoritmico ("auditable by construction")? 8) Regole di pubblicità: quando un audit è negativo, e i problemi non sono ancora risolti, quale dovrebbe essere il comportamento degli auditor, in che modo dovrebbe essere reso pubblico il fallimento di un audit? 9) Chi controlla i controllori? Come assicurarsi che siano davvero responsabili delle proprie azioni?

Un esempio cruciale, emerso dai risultati della nostra ricerca, è il riconoscimento facciale. I sistemi di ADM che sono basati su tecnologie biometriche, incluso il riconoscimento facciale, si sono rivelati una minaccia particolarmente seria per l'interesse pubblico e i diritti umani, visto che aprono la strada a forme di sorveglianza di massa indiscriminata — in particolare modo quando, come ora, vengono adottati perlopiù in modo non trasparente.

Chiediamo che gli usi pubblici di tecnologie di riconoscimento facciale capaci di sorveglianza di massa siano strettamente, e urgentemente, proibiti fino a data da destinarsi, a livello europeo.

Tali tecnologie potrebbero perfino essere considerate già illegali nell'UE, almeno per certi usi, se impiegate senza "consenso specifico" dei soggetti inquadrati. Questa lettura legale è stata suggerita dalle autorità in Belgio, quando hanno emesso una prima, storica multa per l'indebita adozione di un sistema di riconoscimento facciale nel paese.

3. Accrescere l'alfabetismo algoritmico e rafforzare il dibattito pubblico sui sistemi di ADM

Maggiore trasparenza sui sistemi di ADM può essere realmente utile solo se chi la deve realmente affrontare — legislatori, governi, organi di settore — è in grado di confrontarsi con l'impatto di tali sistemi in modo responsabile e prudente. Inoltre, coloro i quali subiscono quell'impatto devono poter essere in grado di comprendere quando, perché e come quei sistemi siano stati adottati. È per questa ragione che dobbiamo accrescere l'alfabetismo algoritmico, a tutti i livelli, sia tra i più importanti portatori di interesse che nel pubblico generalista, e rafforzare i dibattiti pubblici sui sistemi di ADM e il loro impatto sulla società, rendendoli maggiormente plurali.

/ Creare centri di competenze sull'ADM indipendenti

Insieme alle nostre richieste sull'auditing algoritmico e il supporto alla ricerca, chiediamo vengano costruiti dei centri di competenze sull'ADM indipendenti, a livello nazionale, per monitorare, valutare e condurre ricerca sui sistemi di ADM. Tali centri dovrebbero insieme fornire consigli al governo e al settore privato, in coordinamento con legislatori, società civile e accademia, circa le conseguenze per la società e i diritti umani derivanti dall'uso di tali sistemi. Il ruolo complessivo di questi centri è la creazione di un sistema funzionante di *accountability*, oltre alla formazione di competenze specifiche.

I centri nazionali di competenze dovrebbero coinvolgere le organizzazioni della società civile, i gruppi dei portatori di interesse, e gli organismi di *enforcement* attualmente esistenti, tra cui le Autorità per la protezione dei dati personali (DPAs) e gli organismi nazionali deputati al rispetto dei diritti umani, così da arrecare il massimo beneficio all'intero ecosistema e promuovere fiducia, trasparenza, e cooperazione tra tutte le parti in causa.

In qualità di organismi indipendenti, questi centri di expertise avrebbero un ruolo centrale nel coordinare gli sviluppi di policy e le strategie nazionali correlate all'ADM, così come nel contribuire all'ampliamento delle competenze e abilità esistenti tra legislatori, governi e organi di settore, in risposta all'aumentato utilizzo dei sistemi di ADM.

Tali centri non dovrebbero disporre di poteri regolatori, ma piuttosto fornire le competenze necessarie a proteggere i diritti umani degli individui, e impedire danni sociali e alla collettività. Dovrebbero, per esempio, aiutare le piccole e medie imprese a soddisfare i loro obblighi in termini di "due diligence" sui diritti umani, incluse le valutazioni sui diritti umani e di impatto algoritmico, e l'iscrizione al registro pubblico dei sistemi di ADM sopra discusso.

/ Promuovere un dibattito democratico plurale e inclusivo sui sistemi di ADM

Oltre a rafforzare abilità e competenze di chi adotta i sistemi di ADM, è altrettanto vitale promuovere l'alfabetismo algoritmico nel pubblico generalista attraverso un più ampio dibattito e programmi pluralisti.

I nostri risultati suggeriscono non solo che i sistemi di ADM restino opachi al pubblico quando sono già utilizzati, ma che perfino la decisione di adottare o meno un sistema di ADM venga di norma presa a insaputa e senza alcuna partecipazione della cittadinanza.

C'è dunque un urgente bisogno di includere il pubblico (e l'interesse pubblico) nei processi decisionali sui sistemi di ADM fin dal principio.

Più in generale, c'è bisogno di un dibattito pubblico maggiormente plurale sugli impatti dell'ADM. Dobbiamo andare oltre il semplice relazionarsi a gruppi di esperti, e garantire

invece che la questione risulti più accessibile a un pubblico più ampio. Ciò significa parlare una lingua diversa da quella meramente tecno-giudiziaria, ingaggiare il pubblico, e stimolarne la curiosità.

Per riuscirci, dovrebbero essere predisposti programmi dettagliati per fare e promuovere alfabetismo digitale. Se l'obiettivo è un dibattito pubblico informato per la creazione di vera autonomia digitale per i cittadini europei, dobbiamo cominciare a costruirlo, e promuoverlo davvero, con una attenzione specifica alle conseguenze sociali, etiche e politiche dell'adozione dii sistemi di ADM.

UNIONE EUROPEA Alle radici del futuro dell'ADM in Europa

Se i sistemi di decision-making automatizzato (ADM) hanno assunto un ruolo crescente nella distribuzione di diritti e servizi in Europa, anche le istituzioni in tutto il Continente hanno preso a riconoscerne sempre più la presenza nella vita pubblica, sia in termini di opportunità che di sfide.

Di Kristina Penner e Fabio Chiusi

Dal nostro primo rapporto nel gennaio 2019 — e nonostante l'Unione Europea sia ancora indaffarata nel più ampio dibattito su un'intelligenza artificiale "degna di fiducia" ("Trustworthy AI") — diversi organi, dal Parlamento UE al Consiglio d'Europa, hanno pubblicato documenti il cui obiettivo è mettere l'UE e l'Europa tutta nelle condizioni di affrontare l'ADM negli anni, se non nelle decadi, a venire.

Nell'estate del 2019 la neoeletta presidente della Commissione, Ursula Von der Leyen, una "ottimista tecnologica" dichiarata, si è impegnata a proporre "un pacchetto legislativo per un approccio europeo coordinato sulle implicazioni umane ed etiche dell'intelligenza artificiale", e a "regolare l'AI" entro i primi cento giorni in carica. Invece, a febbraio 2020, la Commissione Europea ha pubblicato un 'Libro bianco' sull'AI ('White Paper On Artificial Intelligence - A European approach to excellence and trust') contenente "idee e azioni" — ovvero, un pacchetto strategico che ha l'obiettivo di informare i cittadini e aprire la strada per futuri interventi legislativi. Il documento argomenta anche in favore di una "sovranità tecnologica" europea: nelle parole della stessa Von der Leyen, ciò si traduce nella "capacità che l'Europa deve possedere di fare le proprie scelte, basate sui propri valori, rispettando le proprie regole", e dovrebbe inoltre "contribuire a renderci tutti tecno-ottimisti".

Un secondo fondamentale progetto è significativo per l'ADM in Europa: il Digital Services Act (DSA) annunciato nell"Agenda per l'Europa' di Von der Leyen, e concepito per rimpiazzare la direttiva E-Commerce in vigore fin dal 2000. Il suo scopo è "aggiornare le nostre regole di responsabilità e sicurezza per le piattaforme, i servizi e i prodotti digitali, e completare il nostro Mercato Unico Digitale" — alimentando così un fondamentale dibattito sul ruolo dell'ADM nelle politiche di moderazione dei contenuti, in quelle sulla responsabilità degli intermediari, e per la libertà di espressione[13] più in genere.

Una problematizzazione esplicita dei sistemi di ADM è reperibile nella Risoluzione approvata dalla Commissione sul mercato interno e la protezione dei consumatori del Parlamento UE, e in una Raccomandazione "sugli impatti in termini di diritti umani dei sistemi algoritmici" del Comitato dei Ministri del Consiglio d'Europa.

Il Consiglio d'Europa (CoE), in particolare, ha mostrato nel corso dell'ultimo anno di ricoprire un ruolo sempre più importante nel dibattito su come governare l'AI, e anche se il suo reale impatto sulle proposte di regolamentazione resta da dimostrare, si potrebbe sostenere che funga da "guardiano" dei diritti umani. Ciò è maggiormente evidente nella Raccomandazione, 'Unboxing Artificial Intelligence: 10 steps to protect Human Rights', della Commissaria per i diritti umani del CoE, Dunja Mijatović, e nei lavori dell'Ad Hoc Committe on AI (CAHAI) creato nel settembre 2019.

13 Considerazioni dettagliate e raccomandazioni circa l'uso di sistemi di ADM nel contesto del DSA possono essere reperite tra i risultati di un altro progetto di AlgorithmWatch, 'Governing Platforms', https://algorithmwatch.org/en/project/governing-platforms/

DIVERSI OSSERVATORI VEDONO UNA FONDAMENTALE TENSIONE TRA GLI IMPERATIVI AFFARISTICI E QUELLI DEI DIRITTI NEL MODO IN CUI LE ISTITUZIONI EUROPEE, E IN PARTICOLAR MODO LA COMMISSIONE, STANNO CONCEPENDO LE LORO RIFLESSIONI E PROPOSTE SU AI E ADM.

Diversi osservatori vedono una fondamentale tensione tra gli imperativi affaristici e quelli dei diritti nel modo in cui le istituzioni europee, e in particolar modo la Commissione, stanno concependo le loro riflessioni e proposte su AI e ADM.

Da un lato, l'Europa vorrebbe "incrementare l'uso, e la domanda, di dati e prodotti e servizi resi possibili dai dati in tutto il Mercato Unico", diventando così "leader" nelle applicazioni commerciali dell'AI, e potenziando la competitività delle aziende UE a fronte delle crescenti pressioni dai rivali negli Stati Uniti e in Cina. Tutto questo è ancora più significativo per l'ADM, visto che l'assunto di fondo è che, attraverso un'economia "*data-agile*", l'UE "possa diventare un modello e una guida per una società che grazie ai dati riesce a prendere decisioni migliori — negli affari come nel settore pubblico". Come scrive il White Paper sull'AI, "i dati sono la linfa vitale dello sviluppo economico".

Dall'altro lato, tuttavia, l'analisi automatica di dati riguardanti la salute, il lavoro e i diritti sociali di un cittadino può portare a decisioni dagli esiti discriminatori e iniqui. Questo "lato oscuro" degli algoritmi nei processi di decisione è affrontato nel pacchetto di strumenti regolatori (*toolbox*) proposto dall'UE attraverso una serie di principi. Nel caso di sistemi ad alto rischio, dovrebbero essere imposte regole per garantire che i processi decisionali automatici siano compatibili con i diritti umani e con effettive tutele democratiche. Questo è un approccio che le istituzioni UE definiscono "umano-centrico" e insieme unico, perché fondamentalmente opposto a quelli applicati negli Stati Uniti (dove il criterio è il profitto) e in Cina (dove invece lo sono la sicurezza nazionale e la sorveglianza di massa).

Tuttavia, sono emersi dubbi circa le reali possibilità per l'Europa di raggiungere entrambi gli obiettivi allo stesso tempo. Il caso del riconoscimento facciale è quello più comunemente portato a esempio: anche se, come mostrato da questo rapporto, disponiamo ormai di svariate prove che ne testimoniano l'adozione opaca e incontrollata in buona parte dei paesi membri, la Commissione UE non ha reagito rapidamente e in modo deciso per proteggere i diritti dei cittadini europei. Come rivelato in un leaked draft del 'Libro bianco sull'AI' della Commissione, l'UE aveva considerato l'idea di mettere al bando i sistemi di "identificazione biometrica da remoto" nei luoghi pubblici, prima di cambiare idea all'ultimo minuto e preferivi, al contrario, la chiamata a un "ampio dibattito" sulla materia.

Nel frattempo, diversi progetti finanziati dall'UE continuano a promuovere controversi sistemi di ADM per il controllo dei confini, anche dotati di riconoscimento facciale.

Scelte politiche e dibattiti politici

/ La Strategia europea sui dati e il Libro bianco sull'AI

Se da un lato il promesso pacchetto di norme "per un approccio europeo coordinato sulle implicazioni umane ed etiche dell'intelligenza artificiale", annunciato nell"Agenda per l'Europa' di Von der Leyen, non è stato poi effettivamente proposto entro "i primi cento giorni in carica", dall'altro la Commissione UE ha quantomeno pubblicato una serie di documenti che stabiliscono l'insieme di principi e idee che dovrebbero informarlo.

Il 19 febbraio 2020, infatti, sono stati pubblicati insieme la 'Strategia europei sui dati' ('A European Strategy for Data') e il già menzionato 'Libro bianco sull'intelligenza artificiale', i due documenti che illustrano i principi fondamentali dell'approccio strategico dell'Unione Europea all'AI (e ai sistemi di ADM, anche se non esplicitamente menzionati). Tra questi, spiccano l'idea di mettere "al primo posto" le persone (*people first,* inteso come sinonimo di una "tecnologia al servizio delle persone"), la neutralità tecnologica (nessuna tecnologia è buona o cattiva di per sé; ciò, piuttosto, dipende dall'uso che se ne fa) e, naturalmente, la "sovranità tecnologica" e il tecno-ottimismo. Nelle parole di Von der Leyen: "Vogliamo incoraggiare le nostre aziende, i nostri ricercatori, gli innovatori, gli imprenditori, a sviluppare intelligenza artificiale. E vogliamo incoraggiare i cittadini a fidarsi a usarla. Dobbiamo liberarne il potenziale".

"VOGLIAMO INCORAGGIARE LE NOSTRE AZIENDE, I NOSTRI RICERCATORI, GLI INNOVATORI, GLI IMPRENDITORI, A SVILUPPARE INTELLIGENZA ARTIFICIALE. E VOGLIAMO INCORAGGIARE I CITTADINI A FIDARSI A USARLA. DOBBIAMO LIBERARNE IL POTENZIALE."
URSULA VON DER LEYEN

L'idea di fondo è che a nuove tecnologie non debbano accompagnarsi nuovi valori. Il "nuovo mondo digitale" ipotizzato dall'amministrazione Von der Leyen intende infatti rispettare a pieno i diritti umani e civili. "Eccellenza" e "fiducia", evidenziati fin dal titolo stesso del Libro bianco, vengono considerate i due pilastri su cui un modello europeo per l'AI può e deve poggiare, differenziandolo sia da quello statunitense che da quello cinese.

E tuttavia, all'ambizione non corrisponde adeguato livello di dettaglio nel Libro bianco. Per esempio, il White Paper della Commissione illustra un approccio alla regolamentazione dell'AI basato sul rischio, in cui le regole sono proporzionali alla severità dell'impatto di un sistema di "AI" sulle vite dei cittadini. "Per i casi ad alto rischio, come nella sanità, nella sorveglianza (*policing*) e nei trasporti", si legge, "i sistemi di AI dovrebbero essere trasparenti, tracciabili e garantire supervisione umana". Tra le salvaguardie che dovrebbero venire predisposte figurano anche una fase di test e la certificazione degli algoritmi adottati, con l'obiettivo di renderle comuni come per "cosmetici, vetture o giocattoli". Al contrario, "sistemi meno rischiosi" prevedono unicamente l'osservanza di schemi di controllo qualità volontari: "Gli operatori economici coinvolti si vedrebbero di conseguenza riconosciuti un marchio di qualità per le loro applicazioni di AI".

Diversi critici, tuttavia, hanno sottolineato che la definizione stessa di "rischio" nel Libro bianco è insieme circolare e troppo vaga, finendo per consentire a diversi sistemi di ADM altamente impattanti di insinuarsi tra le maglie dello schema di regole proposto[14].

I commenti[15] pervenuti nel corso della consultazione pubblica tenutasi tra febbraio e giugno 2020 evidenziano quanto questa idea sia controversa, con il 42,5% delle risposte a favore di "requisiti obbligatori" per le sole "applicazioni di AI ad alto rischio", e il 30,6% che invece esprimono dubbi al riguardo.

Inoltre, non vi è nel documento descrizione chiara di un meccanismo per fare poi realmente rispettare i requisiti esposti, né quantomeno di un processo che porti a costruirlo davvero.

Le conseguenze sono immediatamente visibili per le tecnologie biometriche, e per il riconoscimento facciale in particolare. Su questo, il Libro bianco propone una distinzione tra "autenticazione" biometrica, considerata non controversa (per esempio, il riconoscimento facciale per sbloccare uno smartphone), e "identificazione" biometrica da remoto (per esempio, per identificare manifestanti in protesta in una pubblica piazza), che può sollevare serie preoccupazioni in termini di privacy e diritti umani.

Solo i casi rientranti in quest'ultima categoria, tuttavia, sarebbero considerati problematici secondo lo schema di regole ipotizzato dall'UE. Eppure il FAQ del Libro bianco afferma: "questa è la forma più intrusiva di riconoscimento facciale, e sarebbe in linea di principio proibita nell'Unione Europea", a meno che la sua adozione non sia materia di "sostanziale interesse pubblico".

Il documento esplicativo aggiunge anche che "consentire il riconoscimento facciale è attualmente l'eccezione", ma i risultati del nostro rapporto lo contraddicono: il riconoscimento facciale sembra piuttosto stare diventando rapidamente la norma. Una versione preliminare del Libro bianco, ottenuta dagli organi di stampa, pareva avere riconosciuto l'urgenza del problema, includendo l'idea di una moratoria dai tre ai cinque anni degli usi dal vivo del riconoscimento facciale in luoghi pubblici, fino a quando — e se — non si fosse trovato un modo per coniugarli con tutele democratiche.

Appena prima della pubblicazione della bozza ufficiale del Libro bianco, perfino la commissaria europea Margrethe Vestager aveva chiesto che tali usi del riconoscimento facciale venissero messi in "pausa".

Tuttavia, "funzionari della Commissione" hanno immediatamente aggiunto che la "pausa" non avrebbe impedito ai governi nazionali di farvi ricorso secondo le regole esistenti. E alla fine, l'ultima versione del Libro bianco non contiene alcuna menzione di una moratoria per il riconoscimento facciale, preferendo al contrario chiedere "un ampio dibat-

14 "Per fare due esempi: VioGén, un sistema di ADM per prevedere casi di violenza di genere, e Ghostwriter, un'applicazione per identificare frodi durante gli esami, sfuggirebbero alla forma più severa di regolamentazione, nonostante presentino rischi enormi" (https://algorithmwatch.org/en/response-european-commission-ai-consultation/)

15 "In totale, sono stati ricevuti 1215 contributi, dei quali 352 per conto di aziende o di organizzazioni e associazioni commerciali, 406 di cittadini (il 92% dell'UE), 152 per conto di istituti accademici e di ricerca, e 73 da autorità pubbliche. La voce della società civile è stata rappresentata da 160 rispondenti (tra cui 9 associazioni per i consumatori, 129 organizzazioni non governative e 22 sindacati); 72 partecipanti hanno risposto identificandosi nella categoria "altro". Commenti sono giunti "da ogni parte del mondo", inclusi paesi come "India, Cina, Giappone, Siria, Iraq, Brasile, Messico, Canada, Stati Uniti e Gran Bretagna". (Dal 'Summary Report' della consultazione pubblica, reperibile da qui: https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence)

tito europeo sulle specifiche condizioni, ammesso esistano, che potrebbero giustificarne" l'uso per scopi di identificazione biometrica dal vivo. Il Libro bianco vi include i principi di giustificazione e proporzionalità, l'esistenza di tutele democratiche, e il rispetto dei diritti umani.

In tutto il documento, i rischi associati alle tecnologie di AI vengono di norma definiti "potenziali", mentre i benefici vengono al contrario dipinti come realissimi e immediatamente realizzabili. Ciò ha portato molti[16], nella comunità dei diritti umani, ad affermare che la narrazione complessiva del Libro bianco contenga una preoccupante inversione di priorità nelle politiche dell'UE, finendo per privilegiare l'obiettivo di una maggiore competitività globale a quello di proteggere i diritti umani.

E ciononostante, i documenti UE sollevano anche alcune questioni fondamentali. Per esempio, l'interoperabilità dei sistemi di AI, e la creazione di una rete di centri di ricerca specializzati in applicazioni dell'AI che ambiscano all'"eccellenza" e alla creazione di competenze adeguate.

L'obiettivo è "attrarre oltre 20 miliardi l'anno di investimenti complessivi in AI a livello UE, nel corso del prossimo decennio".

16 Per esempio Access Now (https://www.accessnow.org/cms/assets/uploads/2020/05/EU-white-paper-consultation_AccessNow_May2020.pdf), AI Now (https://ainowinstitute.org/ai-now-comments-to-eu-whitepaper-on-ai.pdf), EDRi (https://edri.org/our-work/can-the-eu-make-ai-trustworthy-no-but-they-can-make-it-just/) — e AlgotithmWatch (https://algorithmwatch.org/en/response-european-commission-ai-consultation/).

Il Libro bianco sembra anche risentire di un certo determinismo tecnologico. "È essenziale", vi si legge infatti, "che le amministrazioni pubbliche, gli ospedali, i servizi di utilità e trasporti, i supervisori finanziari, e altri settori di pubblico interesse comincino rapidamente ad adottare prodotti e servizi basati sull'AI nelle loro attività. Un'attenzione specifica sarà posta nei settori della sanità e dei trasporti, dove la tecnologia è matura per un'adozione su larga scala".

Tuttavia, non è chiaro se suggerire una frettolosa adozione di soluzioni di ADM in ogni sfera dell'attività umana sia compatibile con gli sforzi ipotizzati dalla stessa Commissione UE per affrontare le sfide strutturali poste dai sistemi di ADM a diritti ed equità.

/ La risoluzione del Parlamento UE sull'ADM e la protezione dei consumatori

Una Risoluzione, approvata dal Parlamento europeo nel febbraio 2020, ha affrontato più nello specifico i sistemi di ADM nel contesto della protezione dei consumatori. La Risoluzione sottolinea correttamente che "sistemi algoritmici complessi e processi di decision-making automatizzato stanno venendo creati a passo spedito", e che "le opportunità e le sfide poste da queste tecnologie sono molteplici e riguardano virtualmente ogni settore". Il testo ricorda inoltra il bisogno di "una valutazione dello schema legislativo attualmente in vigore nell'UE", così da giudicare se "sia in grado di far fronte all'emergere di AI e ADM".

Chiedendo un "approccio comune a livello UE per lo sviluppo di processi di ADM", la Risoluzione dettaglia i diversi requisiti che ogni sistema di decisioni automatizzate dovrebbe possedere per rimanere all'interno della sfera dei valori europei. I consumatori dovrebbero essere "debitamente informati" circa i modi in cui gli algoritmi influenzano le loro vite, e dovrebbero sempre ottenere l'intervento di un essere umano con poteri decisionali, così che le scelte che coinvolgono forme di automazione possano essere controllate e, se necessario, corrette. I consumatori dovrebbero poi venire anche informati "quando i prezzi di beni o servizi sono stati personalizzati sulla base di ADM e profilazione delle abitudini di consumo".

Nel ricordare alla Commissione UE che è necessario un approccio basato sui rischi che sia attentamente costruito, la Risoluzione evidenzia che le tutele predisposte devono tenere in considerazione che i sistemi di ADM "possono evolvere e agire in modi non contemplati quando inizialmente

SE L'AI È DAVVERO UNA RIVOLUZIONE CHE RICHIEDE UN PACCHETTO LEGISLATIVO APPOSITO, I RAPPRESENTANTI DEMOCRATICAMENTE ELETTI VOGLIONO AVERE VOCE IN CAPITOLO.

introdotti sul mercato", e che le responsabilità non sono sempre facilmente attribuibili nel caso in cui la loro adozione risulti dannosa.

La Risoluzione richiama inoltre l'art. 22 del GDPR, quando nota che un essere umano deve sempre poter intervenire nel caso in cui "siano in gioco legittimi interessi pubblici", e che sia sempre un essere umano a dover essere, in ultima analisi, responsabile delle decisioni riguardanti "le professioni mediche, legali e contabili, nonché il settore bancario". In particolare, una "adeguata" valutazione del rischio dovrebbe precedere ogni forma di automazione di servizi professionali.

Infine, la Risoluzione elenca una serie dettagliata di requisiti di qualità e trasparenza per il governo dei dati: tra gli altri, "l'importanza di usare solo dati di elevata qualità e privi di bias, così da migliorare l'output dei sistemi algoritmici e aumentare fiducia e accettazione nel consumatore"; usare "algoritmi spiegabili e privi di bias"; e il bisogno di una "struttura per la revisione" che consenta ai consumatori "di ottenere revisione umana e fare appello alle decisioni automatizzate che siano definitive e permanenti".

/ Prendere l'iniziativa tramite il "diritto di iniziativa" del Parlamento UE

Nel suo discorso inaugurale, Von der Leyen ha espresso il proprio chiaro supporto per un "diritto di iniziativa" per il Parlamento europeo. "Nel caso in cui quest'Aula, agendo con il mandato della maggioranza dei suoi Membri, adotti Risoluzioni che richiedono alla Commissione di formulare proposte di legge", ha affermato Von der Leyen, "mi impegno a rispondere con un atto legislativo nel pieno rispetto dei principi di proporzionalità, sussidiarietà e di "legiferare meglio" (*better law-making*)".

Se l'AI è davvero una rivoluzione che richiede un pacchetto legislativo apposito, in arrivo — pare — nel primo trimestre del 2021, i rappresentanti democraticamente eletti vogliono avere voce in capitolo. Il che, sommato all'intento espresso da Von der Leyen di aumentarne la capacità legislativa, potrebbe perfino produrre ciò che Politico ha chiamato "il momento del Parlamento". A conferma, diverse commissione parlamentari sono già all'opera sulla stesura di diversi rapporti.

Ciascun rapporto indaga uno specifico aspetto dell'automazione nelle politiche pubbliche; e per quanto ciascuno di essi sia concepito per informare la legislazione sull'AI prossima ventura, ciò è in buona parte rilevante anche per l'ADM.

Per esempio, nel suo 'Framework of ethical aspects of artificial intelligence, robotics and related technologies', la Commissione Giuridica del Parlamento UE chiede l'istituzione di una "Agenzia europea per l'AI" e, al tempo stesso, di una rete di autorità nazionali di supervisione, in ogni stato membro, così da assicurare che decisioni automatizzate attinenti la sfera morale siano, e restino, etiche.

In 'Intellectual property rights for the development of artificial intelligence technologies', la stessa commissione illustra la sua visione per il futuro dell'automazione nella proprietà intellettuale. La bozza del rapporto, per esempio, afferma che "i metodi matematici non sono brevettabili, a meno che non costituiscano invenzioni di natura tecnica", sostenendo però poi, in tema di trasparenza algoritmica,

che "il *reverse engineering* non è che un'eccezione alla regola del segreto commerciale".

Il rapporto si spinge fino a chiedersi come proteggere "creazioni tecniche e artistiche generate da AI, così da incoraggiare questa forma di creazione", immaginando che "certe opere generate da AI possano essere considerate equivalenti a opere intellettuali, e dunque protette da copyright".

Infine, in un terzo documento ('Artificial Intelligence and civil liability'), la Commissione dettaglia un "approccio per la gestione del rischio" in termini di responsabilità civile delle tecnologie di AI. Secondo tale approccio, "la parte meglio in grado di controllare e gestire i rischi correlati alla tecnologia è considerata strettamente responsabile, e unico viatico a contenziosi".

Importanti principi riguardanti l'uso di ADM nel sistema penale possono essere reperiti nel rapporto della Commissione Libertà Civili, Giustizia e Affari Interni intitolato 'Artificial Intelligence in criminal law and its use by the police and judicial authorities in criminale matters'. Dopo un elenco dettagliato di reali usi correnti dell'"AI" — in realtà, di sistemi di ADM — da parte delle forze di polizia[17], la Commissione "considera necessario creare un insieme di regole chiaro ed equo per assegnare responsabilità legali per le potenziali conseguenze negative prodotte da queste avanzate tecnologie digitali".

Ne dettaglia poi alcune caratteristiche: nessuna decisione può essere interamente automatizzata[18], gli algoritmi devono essere spiegabili in modi che siano "intelligibili agli utenti", e deve essere condotta una "obbligatoria valutazione di impatto sui diritti fondamentali (...) di ogni sistema di AI in uso dalle forze dell'ordine o dalla giustizia", prima della sua adozione, più un altrettanto obbligatorio "auditing periodico di tutti i sistemi di AI usati dalle forze dell'ordine e dalla giustizia, per testare e valutare i sistemi algoritmici mentre sono in funzione".

Nel rapporto viene anche poi chiesta una moratoria per l'uso di tecnologie di riconoscimento facciale da parte delle forze dell'ordine, quantomeno "finché gli standard tecnici non potranno essere considerati pienamente rispettosi dei diritti fondamentali, i risultati che ne derivano non saranno non discriminatori, e non ci sarà pubblica fiducia nella necessità e proporzionalità dell'adozione di tali tecnologie".

L'obiettivo è di pervenire a una maggiore trasparenza nell'uso dei sistemi di ADM, e insieme suggerire agli stati membri di fornire una "comprensione esaustiva" dei sistemi di AI adottati dalle forze dell'ordine e nel sistema giudiziario, e — sulla scia dell'idea di un "registro pubblico" — dettagliare "le tipologie di strumenti in uso, le tipologie di crimini a cui si applicano, e le aziende i cui strumenti stiano venendo utilizzati".

La Commissione Cultura e Istruzione e quella per le politiche industriali erano a loro volta al lavoro sui rispettivi rapporti, al momento della stesura di questo capitolo.

Tutte queste iniziative hanno portato, il 18 giugno 2020, alla creazione di una Commissione Speciale sull'intelligenza artificiale nell'era digitale ('Special Committee on Artificial Intelligence in a Digital Age', o 'AIDA'). Composta di 33 membri, e con una durata iniziale di dodici mesi, intende analizzare "il futuro impatto" dell'AI sull'economia dell'Unione, e "in particolare sulle competenze, l'impiego, il fintech, l'educazione, la salute, il turismo, l'agricoltura, l'ambiente, la difesa, l'industria, l'energia e l'e-government".

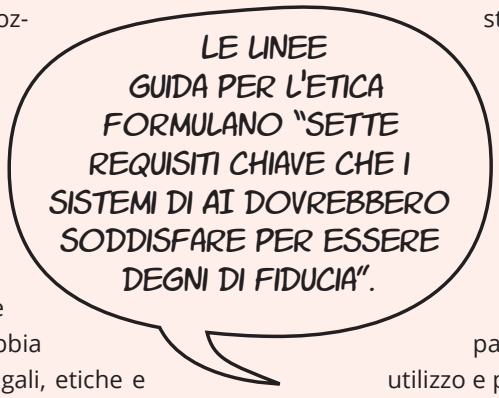
/ L'High-Level Expert Group sull'AI e la AI Alliance

Nel 2018, la Commissione Europea ha predisposto la creazione di un High-Level Expert Group (HLEG) sull'AI, un comitato composto da 52 esperti che ha l'obiettivo di contribuire all'implementazione di una strategia europea sull'intelligenza artificiale, identificare gli obiettivi che dovrebbero venire osservati per ottenere una "AI degna di fiducia" ("*trustworthy AI*") e, in qualità di comitato direttivo della più ampia "Alleanza per l'AI" (*AI Alliance*), creare una piattaforma aperta e multi-stakeholder — composta attualmente da oltre quattromila membri — per fornire più ampi input al lavoro del gruppo di esperti sull'AI.

17 A pagina 5, il rapporto afferma: "Le applicazioni dell'AI usate dalle forze dell'ordine includono: le tecnologie di riconoscimento facciale, il riconoscimento automatico di numeri di targa, l'identificazione di un parlante, l'identificazione di una parlata, tecnologie per la lettura delle labbra, tecnologie di sorveglianza sonora (per esempio, algoritmi di riconoscimento di uno sparo da arma da fuoco), l'autonoma ricerca e analisi nei database identificati, tecnologie predittive (predictive policing e analisi degli hotspot criminali), strumenti di riconoscimento del comportamento, strumenti autonomi per identificare frodi finanziarie e fondi di provenienza terroristica, social media monitorino (scraping e data haversting a caccia di connessioni), IMSI (International Mobile Subscriber Identity) catcher, e strumenti di sorveglianza automatizzata che incorporano diverse funzionalità di riconoscimento (come per esempio i termoscanner e gli strumenti di riconoscimento del battito cardiaco)".

18 "In contesti giudiziari e di polizia, la decisione finale deve sempre essere presa da un essere umano." (p. 6)

Dopo la pubblicazione della prima bozza delle linee guida per l'etica in una intelligenza artificiale "degna di fiducia" ('AI Ethics Guidelines for Trustworthy AI') a dicembre 2018, che ha raccolto feedback da oltre 500 commentatori, ne è stata pubblicata anche una versione aggiornata, ad aprile 2019. Il documento propone un "approccio umano-centrico" che abbia lo scopo di ottenere soluzioni di AI legali, etiche e solide attraverso ogni fase del loro ciclo di vita. Resta, in ogni caso, uno schema volontario senza raccomandazioni concrete e applicabili in termini di operatività, implementazione ed *enforcement*.



La società civile e le organizzazioni per la tutela del consumatore e dei diritti, commentando il documento, hanno chiesto la traduzione delle linee guida in diritti tangibili. Per esempio, la no profit per i diritti digitali Access Now, membro del HLEG, ha insistito affinché il passo successivo della Commissione sia chiarire come i diversi portatori di interesse possano testare, applicare, migliorare, approvare e implementare forme di "trustworthy AI", riconoscendo al contempo il bisogno di determinare i limiti oltre il quale l'approccio europeo non consente di spingersi.

In un op-ed, altri due membri del HLEG hanno affermato che il gruppo aveva "lavorato per un anno e mezzo, solo per poi vedere le proprie proposte perlopiù ignorate o menzionate solo di passaggio" nella bozza finale stilata dalla Commissione Europea. In aggiunta, i due hanno anche scritto che, dato che il gruppo era stato creato con un primo obiettivo di identificare rischi e limiti invalicabili (*red lines*) per l'AI, alcuni suoi membri avevano indicato armi automatiche, sistemi di valutazione a punti per la cittadinanza (*citizen scoring*) e identificazione automatica tramite riconoscimento facciale tra le implementazioni dell'AI da evitare. Tuttavia, soggetti rappresentanti degli interessi dell'industria di settore, dominanti nella composizione del gruppo[19], sono riusciti a ottenere la rimozione di questi principi dalla bozza del documento prima che fosse pubblicata.

Questo sbilanciamento a favore dei potenziali benefici dell'ADM può essere osservato anche nelle pagine del secondo importante documento prodotto dal HLEG. Nel suo rapporto 'Policy and investment recommendations for trustworthy AI in Europe', reso pubblico a giugno 2019, si trovano infatti 33 raccomandazioni mirate a "indirizzare un'AI degna di fiducia nella direzione della sostenibilità, della crescita e della competitività, così come in quella dell'inclusione — rinforzando, beneficiando e proteggendo gli esseri umani". Il documento è principalmente una chiamata a incrementare utilizzo e portata dell'AI nel settore privato come in quello pubblico, investendo in strumenti e applicazioni "che aiutino le fasce più vulnerabili della popolazione", così da "non lasciare indietro nessuno".

Ciononostante, e pur date le legittime perplessità, entrambe le linee guida esprimono anche considerazioni critiche, avanzando alcune richieste per i sistemi di ADM. Per esempio, le linee guida per l'etica formulano "sette requisiti chiave che i sistemi di AI dovrebbero soddisfare per essere *trustworthy*", cioè meritare la fiducia dei cittadini. Non solo: forniscono anche una guida all'implementazione pratica di ogni requisito — autonomia e controllo umano, sicurezza e solidità tecnica, privacy e governo dei dati, trasparenza, diversità, equità e principio di non discriminazione, benessere sociale e ambientale, e responsabilità (*accountability*).

Le linee guida contengono poi anche indicazioni più concrete, in quella che definisce "Trustworthy AI assessment list", un vero e proprio elenco per pervenire all'operatività dei principi appena dettagliati. L'obiettivo è che venga adottato "nello sviluppo, nell'adozione o nell'utilizzo dei sistemi di AI", e adattato "ai casi d'uso specifici in cui il sistema sta venendo applicato".

La lista include diverse questioni associate al rischio di abusare dei diritti umani tramite sistemi di ADM. Tra di esse, figurano la mancanza di autonomia e controllo umano, problemi di natura tecnica e di sicurezza informatica, l'incapacità di evitare bias iniqui o fornire accesso universale a tali sistemi, nonché la mancanza di un reale accesso ai dati che processano.

Contestualmente, questa bozza di elenco inclusa nelle linee guida fornisce suggerimenti utili per chiunque intenda adottare sistemi di ADM. Per esempio, chiede "una valutazione di impatto sui diritti fondamentali quando può esserci un impatto negativo" su quei diritti. Chiede anche se siano stati predisposti "meccanismi specifici di controllo e vigilanza" nei casi riguardanti sistemi di AI "capaci di auto-

19 Il gruppo era composto da 24 rappresentanti del settore commerciale, 17 accademici, 5 organizzazioni della società civile, e 6 ulteriori membri, tra cui l'Agenzia dell'Unione Europea per i diritti fondamentali.

apprendimento o autonomi", e se esistano processi per "assicurare la qualità e l'integrità dei vostri dati".

Considerazioni dettagliate riguardano inoltre alcune questioni fondamentali per i sistemi di ADM, come la loro trasparenza e spiegabilità (*explainability*). Alcune delle domande incluse sono "fino a che punto le decisioni, e dunque i risultati di scelte prese da sistemi di AI, sono comprensibili?" e "fino a che punto una decisione del sistema influenza i processi decisionali dell'organizzazione?" Si tratta di questioni estremamente rilevanti per valutare i rischi posti dalla loro adozione.

In aggiunta, le raccomandazioni sulle politiche di investimento prevedono la determinazione di limiti (*red lines*) attraverso un "dialogo" istituzionalizzato "sulle politiche per l'AI con i portatori di interesse che ne devono fronteggiare l'impatto", inclusi gli esperti della società civile. Sollecitano inoltre a "mettere al bando sistemi massivi di valutazioni a punteggio (*scoring*) degli individui basati su AI, così come definiti nelle linee guida per l'etica, e [a] stabilire regole estremamente chiare e stringenti per la sorveglianza a scopi di sicurezza nazionale, e per altri scopi che siano stati dichiarati di interesse pubblico o nazionale". Il divieto includerebbe anche le tecnologie di identificazione biometrica e di profilazione.

Altro aspetto rilevante per i sistemi di decision-making automatizzato è che il documento afferma anche che "definire chiaramente se, quando e come l'AI può essere usata (...) sarà cruciale per riuscire ad avere una AI davvero degna di fiducia", ammonendo che "qualunque forma di valutazione a punteggio (*scoring*) del cittadino può portare a perdita di autonomia individuale e mettere a repentaglio il principio di non discriminazione", e "di conseguenza tali valutazioni dovrebbero essere usate solo nel caso ve ne sia chiara giustificazione, all'interno di un insieme di misure proporzionate ed eque". Si aggiunge infine che "la trasparenza non basta a prevenire discriminazioni o fare sì che sia garantita l'equità". Ciò significa che la possibilità di rifiutarsi di essere assoggettati a un meccanismo di scoring dovrebbe essere una scelta sempre disponibile, idealmente senza che ne derivi alcuno svantaggio per il cittadino.

Da un lato, il documento riconosce "che, pur arrecando sostanziali benefici all'individuo e alla società, i sistemi di AI pongono anche alcuni rischi, e possono avere impatti negativi, tra cui alcuni che sono difficili da prevedere, identificare o misurare (per esempio, concernenti la democrazia, lo stato di diritto e la giustizia distributiva, nonché la stessa mente umana)". Dall'altro, tuttavia, il gruppo afferma che "regole innecessariamente prescrittive dovrebbero essere evitate".

Nel luglio del 2020, il gruppo sull'AI ha anche introdotto la versione finale della sua lista per la valutazione dell'intelligenza artificiale meritevole di fiducia ('Assessment List for Trustworthy Artificial Intelligence', o 'ALTAI'), ottenuta al termine di un processo deliberativo che ha incluso 350 portatori di interesse.

La lista, la cui adozione è interamente volontaria e priva di ogni implicazione in termini di legge, ha l'obiettivo di tradurre i sette requisiti descritti nelle linee guida per l'etica del HLEG sull'AI in azione. L'intento è fornire uno strumento di auto-valutazione a chiunque intenda costruire soluzioni di AI che siano compatibili con i valori dell'UE — per esempio, progettisti e sviluppatori di sistemi di AI, scienziati dei dati, funzionari e specialisti del *procurement*, funzionari legali e per la *compliance*.

/ Il Consiglio d'Europa: come proteggere i diritti umani nell'ADM

In aggiunta all'Ad Hoc Committee on Artificial Intelligence (CAHAI), creato a settembre 2019, il Comitato dei Ministri del Consiglio d'Europa[20] ha a sua volta reso pubblico un framework dotato di sostanza e profondità.

Immaginata come strumento per la determinazione di standard, la sua 'Raccomandazione agli stati membri sugli impatti dei sistemi algoritmici sui diritti umani' ('Recommendation to Member States on the human rights impacts of algorithmic systems') descrive[21] le "significative sfide" che si presentano in congiunzione con l'emergere di tali sistemi, e con il nostro "crescente affidamento" a essi. Tali sfide

20 Il Consiglio d'Europa (CoE) è insieme "un organo governativo in cui gli approcci nazionali ai problemi europei vengono discussi su base paritaria e un forum di risposte collettive a tali sfide". I suoi ambiti di operatività includono "gli aspetti politici dell'integrazione europea, la tutela delle istituzioni democratiche, dello stato di diritto e dei diritti umani — in altre parole, tutti i problemi che richiedano soluzioni concertate a livello europeo". Anche se le raccomandazioni ai governi degli stati membri non sono vincolanti, in alcuni casi il Comitato dei ministri del CoE può chiedere a quei governi di informarlo circa ogni azione da loro intrapresa in risposta a tali raccomandazioni (Art. 15b dello Statuto). I rapporti tra il CoE e l'UE sono dettagliati 1) nel Compendio di Testi che governa le relazioni tra CoE e UE, e 2) nel Memorandum of Understanding tra CoE e UE, cfr. https://rm.coe.int/1680306052

21 Sotto la supervisione dello Steering Committee on Media and Information Society (CDMSI) e preparato dal Committee of Experts on Human Rights Dimensions of Automated Data Processing and Different Forms of Artificial Intelligence (MSI-AUT).

sono anche definite rilevanti "per le società democratiche e lo stato di diritto".

Il documento, sottoposto a un periodo di consultazione pubblica in cui ha ricevuto commenti dettagliati da diverse organizzazioni della società civile, va ben oltre il Libro bianco della Commissione UE, in tema di salvaguardia di valori e diritti umani.

La Raccomandazione analizza attentamente gli effetti e le sempre nuove configurazioni dei sistemi algoritmici (Appendice A), concentrandosi su tutte le fasi del processo di costruzione di un algoritmo — a dire: il *procurement*, la progettazione, lo sviluppo, la sua adozione e mantenimento.

Pur seguendo in linea generale l'approccio "umano-centrico" all'AI delle linee guida HLEG, il documento delinea "obblighi" operativi "per gli Stati" (Appendice B), così come responsabilità per gli operatori del settore privato (Appendice C). In più, la Raccomandazione vi aggiunge principi quali "l'autodeterminazione informativa"[22], elenca dettagliati suggerimenti per costruire meccanismi di *accountability* e rimedi efficaci, e chiede l'istituzione di valutazioni di impatto sui diritti umani.

Pur riconoscendo chiaramente "il significativo potenziale delle tecnologie digitali di fare fronte alle sfide sociali, produrre innovazione benefica per la società, e sviluppo economico", il documento invita anche nel contempo alla cautela. L'intento è assicurare che quei sistemi non perpetuino, deliberatamente o accidentalmente, "discriminazioni razziali, di genere, e altri disequilibri sociali e nella forza lavoro che ancora non sono stati eliminati dalle nostre società".

Al contrario, i sistemi algoritmici dovrebbero essere usati in modo proattivo e attento per combattere tali disuguaglianze, e prestare "attenzione ai bisogni e alle voci dei gruppi più vulnerabili".

Anche più significativo, tuttavia, è che la Raccomandazione identifichi rischi potenzialmente più elevati per i diritti umani nel caso in cui siano gli stati membri a usare sistemi algoritmici per fornire servizi pubblici. Dato che è impossibile per un cittadino sottrarvisi, quantomeno senza che ciò comporti conseguenze negative, c'è bisogno di precauzioni e tutele per l'uso di sistemi di ADM nel governo e nell'amministrazione pubblica .

La Raccomandazione affronta anche i conflitti e le sfide derivanti dalle partnership tra pubblico e privato ("né chiaramente pubbliche né chiaramente private") in un ampio spettro di utilizzi.

Le raccomandazioni per i governi degli stati membri includono l'abbandono e il rifiuto di processi e sistemi di ADM nel caso in cui "il controllo e la supervisione umana diventino impraticabili", o quando mettano a repentaglio i diritti umani; e l'adozione di sistemi di ADM se e solo se trasparenza, responsabilità, legalità, e la protezione dei diritti umani possono essere garantite "in ogni fase del processo". Inoltre, il controllo e la valutazione di tali sistemi dovrebbero essere "costanti", "inclusive e trasparenti", ricavate da una interlocuzione con tutti i portatori di interesse rilevanti, così come da un'analisi dell'impatto ambientale e di altre potenziali esternalità su "popolazioni e ambienti".

Nell'appendice A, il Consiglio d'Europa fornisce anche una definizione di cosa intenda per sistemi algoritmici "ad alto rischio" che potrebbe essere d'ispirazione per altri organi e istituzioni. Più nello specifico, la Raccomandazione stabilisce che "il termine "ad alto rischio" si applica in riferimento all'uso di sistemi algoritmici in processi e decisioni che possono produrre serie conseguenze individuali, o in situazioni in cui la mancanza di alternative comporti una probabilità particolarmente elevata di violazione dei diritti umani, incluso tramite l'introduzione o l'amplificazione di ingiustizie distributive".

Il documento, che non ha richiesto approvazione all'unanimità, non è vincolante.

/ Regole per i contenuti terroristici online

Dopo un lungo periodo di apatia, la regolamentazione per prevenire la diffusione di contenuti terroristici ha trovato un nuovo slancio nel corso del 2020. Se dovesse ancora includere la previsione di strumenti automatizzati e proattivi per identificare e rimuovere contenuti online, questi molto probabilmente ricadrebbero tra i casi regolati dall'art. 22 del GDPR.

Come scrive il Garante europeo per la protezione dei dati personali (European Data Protection Supervisor, o 'EDPS'): "dato che gli strumenti automatizzati, così come predisposti nella Proposta, possono portare non solo alla rimozione e

22 "Gli Stati dovrebbero assicurare che il design, lo sviluppo e il mantenimento durante l'operatività dei sistemi algoritmici forniscano agli utenti modalità per essere preventivamente informati circa il trattamento dei dati a essi correlati (inclusi gli scopi e i possibili risultati), nonché di avere il controllo dei loro dati, incluso attraverso forme di interoperabilità", si legge nella Sezione 2.1 dell'Appendice B.

alla conservazione di contenuti (e relativi dati) concernenti l'uploader, ma anche, in ultima analisi, a indagini penali, tali strumenti possono avere significative conseguenze a livello personale, impattando sul diritto alla libera espressione e ponendo una rilevante sfida al godimento di diritti e libertà". Di conseguenza, per questi strumenti si applica l'art. 22(2).

Inoltre, ed è cruciale, il pacchetto normativo richiederebbe tutele ben più sostanziali di quelle attualmente previste dalla Commissione. Come spiega l'organizzazione per i diritti digitali European Digital Rigths (EDRi), "la proposta di regolamentazione per i contenuti terroristici richiede significativi aggiustamenti per essere all'altezza dei valori dell'Unione, e proteggere i diritti fondamentali e le libertà dei cittadini".

Una prima serie di forti critiche alla proposta iniziale è venuta da gruppi della società civile, commissioni del Parlamento europeo, incluse opinioni e analisi dell'Agenzia UE per i diritti fondamentali e la stessa EDRi, così come da un rapporto co-firmato da tre Special Rapporteurs delle Nazioni Unite. Insieme, ne hanno sottolineato le possibili minacce in termini di libertà di espressione e informazione, libertà e pluralismo dei media, libertà di iniziativa economica e protezione dei dati personali.

Tra gli aspetti critici, figurano una insufficiente definizione di "contenuto terroristico", la portata della regolamentazione (attualmente, sono inclusi contenuti a scopo educativo e giornalistico), le "misure proattive" sopra menzionate, una mancanza di controllo giudiziario, obblighi di trasparenza insufficienti per le forze dell'ordine, e l'assenza di tutele "in casi in cui c'è ragione di ritenere che vi sia un impatto sui diritti umani" (EDRi 2019).

L'EDPS sottolinea che tali "adeguate tutele" dovrebbero includere il diritto di ottenere un intervento umano e il diritto alla spiegazione delle decisioni ottenute tramite mezzi automatizzati (EDRi 2019).

Sebbene le tutele suggerite o richieste abbiano trovato poi spazio nella bozza di rapporto del Parlamento europeo sulla proposta, resta ancora da stabilire chi avrà l'ultima parola, in vista del voto finale. Durante i trilogghi a porte chiuse tra Parlamento, Commissione e Consiglio europeo (cominciati a ottobre 2019), un documento fuoriuscito (*leaked*) sostiene che siano possibili unicamente aggiustamenti di lieve entità.

Supervisione e regolamentazione

/ Prime decisioni sul rispetto del GDPR da parte di sistemi di ADM

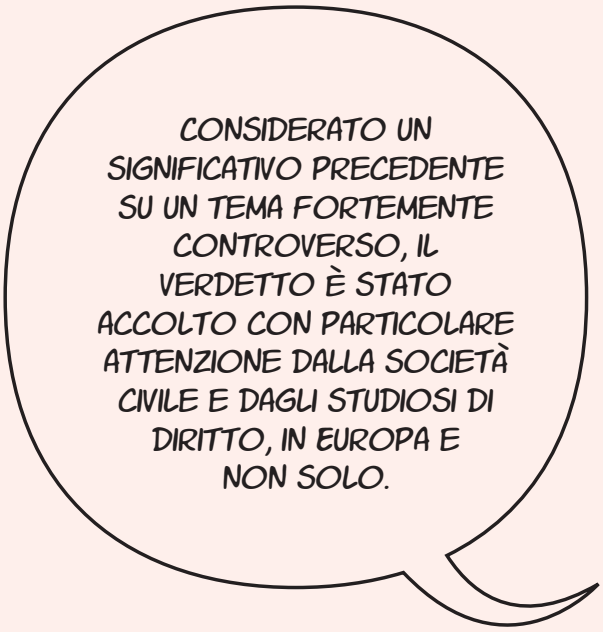
"Sebbene non ci sia stato un ampio dibattito sul riconoscimento facciale durante il passaggio delle negoziazioni sul GDPR e sulla direttiva per la protezione dei dati personali in uso dalle forze dell'ordine, le norme sono state costruite in modo tale da adattarsi nel tempo all'evolvere delle tecnologie. [...] È ora, mentre discute l'etica dell'AI e il suo bisogno di regole, che l'Europa deve determinare se — e se mai — le tecnologie di riconoscimento facciale possano essere consentite in una società democratica. Solo se la risposta è affermativa, e solo allora, potremo rivolgerci alla domanda su quali tutele e sistemi di accountability predisporre" — Wojciech Wiewiórowski, Garante europeo della protezione dei dati

"Il riconoscimento facciale è un meccanismo biometrico particolarmente intrusivo, che comporta seri rischi di abusi della privacy e delle libertà civili delle persone coinvolte" — (CNIL 2019)

Rispetto al precedente rapporto del progetto 'Automating Society', abbiamo testimoniato i primi casi di sentenze e multe correlate a violazioni delle regole emanate dalle Autorità per la protezione dei dati (DPA) nazionali sulla base del GDPR. I casi di studio che stiamo per illustrare, tuttavia, mostrano anche i limiti di applicazione pratica del GDPR in relazione all'art. 22 sui sistemi di ADM, e come stiano costringendo i Garanti nazionali a emettere giudizi caso per caso.

In Svezia, per esempio, un progetto di riconoscimento facciale in fase di sperimentazione in una classe scolastica per un limitato periodo di tempo è stato ritenuto in violazione di diversi requisiti delle norme sulla protezione dei dati (in particolare modo gli artt. 2(14) e 9(2) del GDPR). (European Data Protection Board 2019)

Un tentativo simile risulta attualmente sospeso dopo che la Commission Nationale de l'Informatique et des Libertés (CNIL) francese ha sollevato le proprie preoccupazioni circa l'idea, di due licei, di introdurre tecnologia di riconoscimento facciale in partnership con il colosso statunitense Cisco.



Il suo giudizio tuttavia non è vincolante, e l'azione legale che ne è risultata è ancora in corso[23].

Non è richiesta alcuna previa autorizzazione da parte delle Authority affinché sia possibile condurre simili sperimentazioni, dato che il consenso degli utenti è comunemente considerato sufficiente per processare dati biometrici. Eppure, nel caso svedese, non lo è stato, per via del mancato bilanciamento di poteri tra il *data controller*, titolare del trattamento dei dati, e il *data subject*, o interessato. Sono state dunque ritenute necessarie un'adeguata valutazione di impatto e una previa consultazione con la DPA nazionale.

Lo ha confermato anche l'EDPS:

"Il consenso dovrebbe essere esplicito e dato liberamente, informato e specifico. Eppure non c'è dubbio che una persona non possa esercitare né un opt out, né tantomeno un opt in, qualora necessiti di accesso a luoghi pubblici che sono monitorati da tecnologie di riconoscimento facciale. [...] Infine, la rispondenza di questa tecnologia a principi come la minimizzazione dei dati e la protezione dei dati by design è fortemente in dubbio. Il riconoscimento facciale non è mai stato completamente accurato, e ciò ha serie conseguenze per gli individui erroneamente identificati come criminali o altro. [...] Sarebbe un errore, tuttavia, concentrarsi solo sulle questioni riguardanti la privacy. Questa in una società democratica è invece, al fondamento, una questione etica." (EDPS 2019)

Access Now ha commentato:

"Ora che viene sviluppato un numero crescente di progetti di riconoscimento facciale, osserviamo che il GDPR già fornisce utili tutele in termini di diritti umani, che possono essere richiamate qualora vi siano raccolta e utilizzo illeciti di dati sensibili, per esempio biometrici. Ma le aspettative irresponsabili e spesso infondate createsi intorno all'efficienza di queste tecnologie, oltre agli interessi economici che le motivano, potrebbero condurre a tentativi, da parte di governi centrali e locali, di aggirare la legge".

/ Il riconoscimento facciale usato dalla polizia del Galles del Sud è illegale

Nel corso del 2020, la Gran Bretagna ha testimoniato una prima importante applicazione della Law Enforcement Directive (2016/80, o 'Direttiva Polizia'[24]) circa l'uso di tecnologie di riconoscimento facciale in spazi pubblici da parte delle forze dell'ordine. Considerato un significativo precedente su un tema fortemente controverso, il verdetto è stato accolto con particolare attenzione dalla società civile e dagli studiosi di diritto, in Europa e non solo [25].

Il caso è stato portato in tribunale da Ed Bridges, un 37enne di Cardiff, che aveva affermato che il suo volte fosse stato scannerizzato senza il proprio consenso sia durante lo shopping natalizio del 2017, sia a una protesta pacifica contro le armi da fuoco un anno dopo.

La corte aveva in prima battuta confermato l'uso di "tecnologie di riconoscimento facciale automatico" (*automated facial recognition*, o "AFR") da parte della polizia del Galles del Sud, dichiarandolo legale e proporzionato. Ma il verdetto è stato appellato da Liberty, un gruppo per i diritti civili, e la Corte d'Appello d'Inghilterra e Galles ha rovesciato il rigetto dell'Alta Corte, dichiarando l'uso di AFR illegale l'11 agosto 2020.

23 Si veda il capitolo sulla Francia sull'edizione integrale del rapporto, e (Kaylaki 2019).

24 La 'Direttiva polizia' ('Law Enforcement Directive'), in vigore dal maggio 2018, "concerne il trattamento di dati personali per scopi di polizia — che non ricadono all'interno del campo di applicazione del GDPR" (https://www.dataprotection.ie/en/organisations/law-enforcement-directive)

25 La sentenza è stata emessa il 4 settembre 2019 dall'Alta Corte riunita a Cardiff nel caso Bridges v. the South Wales Politce (High Court of Justice 2019)

Nel suo verdetto, contrario alla polizia del Galles del Sud per tre dei cinque capi di imputazione (*grounds*), la Corte d'Appello giudica infatti di avere trovato "fondamentali lacune" nello schema normativo vigente sull'uso di AFR, concludendo che la sua adozione non soddisfasse il principio di proporzionalità e, inoltre, che non fosse stata condotta una adeguata valutazione di impatto in termini di protezione dei dati (Data Protection Impact Assessment, DPIA), risultandone mancanti diversi passaggi.

La corte non ha, tuttavia, stabilito che il sistema stesse producendo risultati discriminatori, sulla base di genere o razza, dato che la polizia non aveva raccolto prove sufficienti a formulare un giudizio al riguardo[26]. Ciononostante, la Corte ha sentito il bisogno di aggiungere una considerazione degna di nota: "Dato che il riconoscimento facciale automatico è una tecnologia nuova e controversa, ci aspettiamo che tutte le forze di polizia che intendono usarla in futuro si assicurino di avere fatto tutto ciò che è ragionevolmente possibile per garantire che il software non contenga bias razziali o di genere".

Dopo il verdetto, Liberty ha chiesto alla polizia del Galles del Sud e ad altre forze di polizia di abbandonare l'uso di AFR.

ADM in pratica: gestione e controllo dei confini

Mentre la Commissione UE e i relativi portatori di interesse erano impegnati a dibattere se regolare o mettere al bando il riconoscimento facciale, svariate sperimentazioni ne erano già in corso in tutta Europa.

Questa sezione analizza il legame, cruciale e spesso trascurato, tra tecnologie biometriche e sistemi di gestione dei confini dell'UE, mostrando chiaramente come tecnologie che possono produrre risultati discriminatori vengano ugualmente applicate agli individui — per esempio, i migranti — che già scontano maggiormente le discriminazioni esistenti.

/ Riconoscimento facciale e uso di dati biometrici in politiche e prassi dell'UE

Nel corso dell'ultimo anno, il riconoscimento facciale e altre tipologie di tecnologie di identificazione biometrica hanno suscitato crescenti attenzioni da parte di governi, dell'UE, della società civile, e di organizzazioni a difesa dei diritti, in particolare modo riguardo al loro uso da parte delle forze dell'ordine e per la gestione dei confini.

Nel 2019, la Commissione europea ha dato a un consorzio di agenzie pubbliche il compito di "pervenire a una mappa della situazione attuale circa l'uso di tecnologie di riconoscimento facciale in indagini penali in tutti gli stati membri", così da procedere "verso un possibile scambio di dati facciali". Alla società di consulenza Deloitte è stato dunque chiesto di condurre uno studio di fattibilità per un'espansione del sistema Prüm di immagini facciali. Prüm è un sistema applicato a livello europeo che connette tra loro DNA, impronte digitali, e database di immatricolazione dei veicoli, così da consentire ricerche incrociate. La preoccupazione è che un database di volti a livello europeo possa condurre a forme di sorveglianza pervasiva, ingiustificata o illegale.

/ Sistemi di gestione dei confini senza confini

Come riportato nell'edizione precedente di questo rapporto, l'implementazione del sistema complessivo, interoperabile e "intelligente" di gestione dei confini UE, inizialmente proposto dalla Commissione già nel 2013, è in dirittura d'arrivo. Sebbene i nuovi sistemi annunciati (EES, ETIAS[27], ECRIS-TCN[28]) non saranno operativi che nel 2022, la regolamentazione sul sistema ESS (Entry/Exit System) ha già intro-

26 La polizia ha affermato di non avere avuto accesso alla composizione demografica dei dataset utilizzati per il training dell'algoritmo utilizzato, "Neoface". La Corte nota che "resta il fatto, tuttavia, che la SWP (la polizia del Galles del Sud, ndr) non ha mai tentato di verificare da sé, né direttamente né attraverso verifica indipendente, che il software adottato nel caso in esame non contenesse inaccettabili distorsioni in termini razziali o di genere".

27 ETIAS (EU Travel Information and Authorisation System), è il nuovo sistema per la gestione dei confini UE sviluppato da eu-LISA per cittadini "*visa-waiver*", cioè che non necessitano di visto all'ingresso. "Le informazioni immesse durante il processo di applicazione saranno incrociate automaticamente con i database UE esistenti (Eurodac, SIS e VIS), i sistemi futuri EES e ECRIS-TCN, e i database Interpol rilevanti. Ciò renderà possibili sistemi avanzati di verifica di rischi alla sicurezza, alla salute pubblica e di immigrazione irregolare". (ETIAS 2019)

28 L'European Criminal Records Information System — Third Country Nationals (ECRIS-TCN), che dovrà essere sviluppato da eu-LISA, sarà un sistema centralizzato per integrare l'attuale database UE di dati di rilevanza penale (ECRIS) con informazioni su cittadini non-UE condannati nell'Unione Europea.

dotto il concetto delle immagini facciali come identificativi biometrici, nonché, per la prima volta nell'impianto normativo europeo, l'uso di tecnologie di riconoscimento facciale per scopi di autenticazione (*verification*)[29].

L'Agenzia UE per i diritti fondamentali (European Fundamental Rights Agency, o "FRA") ha confermato le modifiche: "ci si attende un'introduzione più sistematica del trattamento di immagini facciali nei sistemi adottati su larga scala a livello UE in tema di asilo, migrazioni, e sicurezza [...] una volta che i necessari passaggi legali e tecnici siano stati completati".

Secondo Ana Maria Ruginis Andrei, di eu-LISA (European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice), questa nuova e ampliata infrastruttura di interoperabilità è stata "assemblata in modo da ottenere il motore adatto a combattere con successo le minacce alla sicurezza interna, controllare efficacemente i flussi migratori e superare gli attuali punti ciechi concernenti la gestione dell'identità". In pratica, ciò significa "detenere le impronte digitali, le immagini facciali, e altri dati personali di fino a 300 milioni di cittadini di paesi non appartenenti all'UE, sommando i dati da cinque sistemi diversi". (Campbell 2020)

/ ETIAS: controlli automatizzati di sicurezza al confine

L'European Travel Information and Authorization System (ETIAS), non ancora operativo mentre scriviamo, sarà condiviso da diversi database per automatizzare i controlli di sicurezza di viaggiatori non appartenenti all'UE (e che non necessitano di un visto) prima che arrivino in Europa.

Il sistema raccoglierà e analizzerà insiemi di dati per ottenere una avanzata "verifica di potenziali rischi correlati alla sicurezza o a migranti irregolari". (ETIAS 2020) Gli obiettivi sono "facilitare i controlli al confine; evitare burocrazia e ritardi per i viaggiatori che si presentano ai confini; assicurare una valutazione coordinata ed equilibrata dei rischi presentati da cittadini di paesi terzi". (ETIAS 2000)

Ann-Charlotte Nygård, a capo dell'unità di "Technical Assistance and Capacity Building" della FRA, vede due specifici rischi concernenti il sistema ETIAS: "il primo è l'uso di dati che potrebbero portare alla discriminazione involontaria di certi gruppi, per esempio se un candidato proviene da una particolare etnia ad alto rischio di migrazione; il secondo riguarda valutazioni dei rischi in termini di sicurezza che siano ricavate da condanne pregresse nel proprio paese d'origine. Alcune di queste condanne pregresse, infatti, potrebbero essere considerate irragionevoli in Europa, come per esempio quelle, in alcuni paesi, a persone di orientamento LGBT. Per evitarlo, [...] gli algoritmi devono essere sottoposti ad auditing, garantendo così che non discriminino; un auditing di questo tipo dovrebbe anche coinvolgere esperti di diverse discipline". (Nygård 2019)

/ iBorderCtrl: riconoscimento facciale e *risk scoring* ai confini

iBorderCtrl è stato un progetto che coinvolgeva le agenzie di sicurezza di Ungheria, Lettonia e Grecia, e che aveva l'obiettivo di "consentire controlli ai confini più rapidi ed esaustivi per cittadini di paesi terzi che abbiano intenzione di attraversare un confine di terra di un paese membro dell'UE". iBorderCtrl faceva ricorso a tecnologie di riconoscimento facciale, una macchina della verità, e a un sistema di valutazione a punteggio (*scoring system*) per sollecitare l'intervento di un agente in carne e ossa nel caso in cui il sistema avesse concluso che un certo individuo è pericoloso, oppure che il suo diritto di ingresso nel paese di destinazione è discutibile.

Il progetto si è concluso nell'agosto del 2019 con risultati contraddittori, nell'ottica di una potenziale implementazione in tutta l'UE.

Sebbene "resti da definirsi l'utilizzo futuro del sistema o di sue parti", la pagina dei "risultati" (*outcomes*) del progetto intravede "la possibilità di integrare funzionalità simili nel nuovo sistema ETIAS, e di estendere la possibilità di trasferire le procedure per i controlli al confine direttamente lì dove si trovano i viaggiatori (autobus, auto, treni, etc.)".

Tuttavia, i moduli a cui ciò si riferisce non sono stati specificati, e i correlati strumenti di ADM sottoposti a test non sono stati analizzati pubblicamente.

Allo stesso tempo la FAQ del progetto conferma che il sistema sperimentato non è da considerarsi "attualmente adatto all'adozione ai confini (...) a causa della sua natura di prototipo e dell'infrastruttura tecnologica esistente a livello UE". Ciò significa che "per un uso da parte delle autorità ai confini sarebbero richiesti futuri sviluppi e un'integrazione al sistema europeo vigente".

29 L'ESS entrerà in funzione nel primo trimestre del 2022, ed ETIAS vi farà seguito entro la fine del 2022. Entrambi sono concepiti come "game changer nel settore della giustizia europea e degli affari interni".

In particolare, mentre l'iBorderCtrl Consortium è stato in grado di dimostrare, in linea di principio, la funzionalità di tali tecnologie per i controlli ai confini, è altrettanto chiaro che questioni di natura etica, legale e sociale debbano ancora venire affrontate prima di qualunque reale adozione.

/ Progetti collegati a Horizon 2020

Diversi progetti, dedicati alla sperimentazione e allo sviluppo di nuovi sistemi e tecnologie per la gestione e il controllo dei confini, vi hanno fatto seguito nell'ambito del programma Horizon2020. I progetti sono elencati sul sito di CORDIS, che fornisce informazioni su tutte le attività di ricerca promosse dall'UE e a quest'ultimo collegate.

Il sito mostra che ci sono 38 progetti attualmente in corso nella sezione tematica 'H2020-EU-3.7.3. — Strengthen security through border management', cioè "rafforzare la sicurezza attraverso la gestione dei confini". Il più ampio programma di cui è parte, 'Secure societies — Protecting freedom and security of Europe and its citizens', vanta un budget complessivo di quasi 1,7 miliardi di euro, e finanzia 350 progetti. L'obiettivo è fare fronte all'"insicurezza, derivante da crimini, violenze, terrorismo, disastri naturali o causati dall'uomo, cyber-attacchi o violazioni della privacy, e altre forme di disordini economici e sociali che sempre più affliggono i cittadini" attraverso progetti che in larga parte consistono nello sviluppo di nuovi sistemi tecnologici a base di AI e ADM.

Alcuni progetti si sono già conclusi e/o loro applicazioni sono già in uso: per esempio, FastPass, ABC4EU, MOBILEPASS e EFFISEC. Ciascuno di essi ha indagato i requisiti necessari a ottenere "un controllo automatizzato dei confini (Automated Border Control, o "ABC") che sia integrato e interoperabile", e dunque applicabile ai sistemi di identificazione e ai gate "intelligenti" nei diversi contesti di attraversamento di un confine.

TRESPASS è un progetto ancora in corso, cominciato nel giugno 2018 e il cui termine è previsto a novembre 2021. La UE contribuisce a finanziarlo con quasi otto milioni di euro, e i coordinatori di iBorderCtrl (così come di FLYSEC e XPDITE) mirano a "sfruttare i risultati e i concetti implementati e sperimentati" da IborderCtrl ed "espanderli in una soluzione di sicurezza multimodale basata sul rischio per i controlli al confine, dotata di una solida base legale ed etica". (Horizon2020 2019)

Il progetto ha lo scopo dichiarato di far evolvere i controlli di sicurezza agli attraversamenti dei confini da un sistema vecchio e superato basato sulle regole (*rule-based*), a un nuovo sistema basato sui rischi (*risk-based*). Ciò significa includere l'applicazione di tecnologie biometriche e sensori, di un sistema di gestione basato sul rischio, e l'uso di modelli adeguati a valutare identità, averi, capacità, e intenti. L'obiettivo è consentire controlli attraverso "link a sistemi antiquati e database esterni come VIS/SIS/PNR", oltre a raccogliere dati da tutte le fonti sopra menzionate per scopi di sicurezza.

L'UE VI CONTRIBUISCE CON POCO MENO DI 8,2 MILIONI DI EURO, PER SVILUPPARE "METODI PIÙ PROGREDITI DI SORVEGLIANZA DEI CONFINI" E CONTRASTARE L'IMMIGRAZIONE IRREGOLARE.

Un altro progetto-pilota, FOLDOUT, è cominciato nel settembre 2018 e si concluderà a febbraio 2022. L'UE vi contribuisce con poco meno di 8,2 milioni di euro, per sviluppare "metodi più progrediti di sorveglianza dei confini" e contrastare l'immigrazione irregolare. Il progetto ha un focus specifico sul "riconoscimento di individui anche in mezzo al fogliame più denso e in condizioni climatiche estreme", combinando "diversi sensori e tecnologie, e fondendole intelligentemente in una piattaforma di identificazione intelligente e affidabile" che suggerisca possibili scenari di intervento. Sperimentazioni sono in corso in Bulgaria, con modelli dimostrativi reperibili in Grecia, Finlandia e nella Guyana Francese.

MIRROR, o 'Migration-Related Risks caused by misconceptions of Opportunities and Requirement', è invece un progetto iniziato a giugno 2019 e il cui termine è previsto a maggio 2022. L'UE vi contribuisce con poco più di cinque milioni di euro, con lo scopo di "comprendere come l'Europa sia percepita all'estero, riconoscere discrepanze tra percezione e realtà, individuare esempi di manipolazione dei media, e sviluppare le abilità necessarie a ribattere alle rappresentazioni scorrette e alle minacce in termini di sicurezza che ne derivano". Sulla base di una "analisi dei rischi specificamente dedicata alle percezioni, il progetto MIRROR affiancherà metodi per l'analisi di testi automatizzati, dei social media e di diversi media (*multimedia*) agli studi empirici", così da sviluppare "tecnologie e nozioni operative [che siano] validate attentamente insieme agli organi deputati al controllo dei confini e ai policy-maker, per esempio tramite sperimentazioni".

Tra gli ulteriori progetti che si sono già conclusi, ma che sono ugualmente menzionati, è incluso TABULA RASA ('Trusted Biometrics under Spoofing Attacks'), cominciato nel novembre 2010 e conclusosi nell'aprile 2014. TABULA RASA analizzava "i punti deboli dei software di identificazione biometrica in relazione alla loro vulnerabilità a falsificazione dell'identità (*spoofing*), che ne diminuisce l'efficacia". Tra i progetti è incluso anche Bodega, cominciato a giugno 2015 e terminato nell'ottobre 2018, che cercava di comprendere come utilizzare "l'expertise del fattore umano" nella "introduzione di sistemi più intelligenti di controlli ai confini, come gate automatici e sistemi di sicurezza self-service basati su dati biometrici".

Riferimenti bibliografici

Access Now (2019): Comments on the draft recommendation of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems https://www.accessnow.org/cms/assets/uploads/2019/10/Submission-on-CoE-recommendation-on-the-human-rights-impacts-of-algorithmic-systems-21.pdf

AlgorithmWatch (2020): Our response to the European Commission's consultation on AI https://algorithmwatch.org/en/response-european-commission-ai-consultation/

Campbell, Zach/Jones, Chris (2020): Leaked Reports Show EU Police Are Planning a Pan-European Network of Facial Recognition Databases https://theintercept.com/2020/02/21/eu-facial-recognition-database/

CNIL (2019): French privacy regulator finds facial recognition gates in schools illegal https://www.biometricupdate.com/201910/french-privacy-regulator-finds-facial-recognition-gates-in-schools-illegal

Coeckelbergh, Mark / Metzinger, Thomas(2020): Europe needs more guts when it comes to AI ethics https://background.tagesspiegel.de/digitalisierung/europe-needs-more-guts-when-it-comes-to-ai-ethics

Committee of Ministers (2020): Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154

Commissioner for Human Rights (2020): Unboxing artificial intelligence: 10 steps to protect human rights https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights

Committee on Legal Affairs (2020): Draft Report: With recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies https://www.europarl.europa.eu/doceo/document/JURI-PR-650508_EN.pdf

Committee on Legal Affairs (2020): Artificial Intelligence and Civil Liability https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf

Committee on Legal Affairs (2020): Draft Report: On intellectual property rights for the development of artificial intelligence technologies https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.pdf

Committee on Civil Liberties, Justice and Home Affairs (2020): Draft Report: On artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters https://www.europarl.europa.eu/doceo/document/LIBE-PR-652625_EN.pdf

Delcker, Janosch(2020): Decoded: Drawing the battle lines — Ghost work — Parliament's moment https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-drawing-the-battle-lines-ghost-work-parliaments-moment/?utm_source=POLITICO.EU&utm_campaign=5a7d137f82-EMAIL_CAMPAIGN_2020_09_09_08_59&utm_medium=email&utm_term=0_10959edeb5-5a7d137f82-190607820

Data Protection Commission(2020): Law enforcement directive https://www.dataprotection.ie/en/organisations/law-enforcement-directive

EDRi (2019): FRA and EDPS: Terrorist Content Regulation requires improvement for fundamental rights https://edri.org/our-work/fra-edps-terrorist-content-regulation-fundamental-rights-terreg/

GDPR (Art 22): Automated individual decision-making, including profiling https://gdpr-info.eu/art-22-gdpr/

European Commission (2018): White paper: On Artificial Intelligence - A European approach to excellence and trust https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Commission (2020): A European data strategy https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

European Commission (2020): Shaping Europe's digital future – Questions and Answers https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_264

European Commission (2020): White Paper on Artificial Intelligence: Public consultation towards a European approach for excellence and trust https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence

European Commission (2018): Security Union: A European Travel Information and Authorisation System - Questions & Answers https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_4362

European Data Protection Board (2019): Facial recognition in school renders Sweden's first GDPR fine https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_en

European Parliament (2020): Artificial intelligence: EU must ensure a fair and safe use for consumers https://www.europarl.europa.eu/news/en/press-room/20200120IPR70622/artificial-intelligence-eu-must-ensure-a-fair-and-safe-use-for-consumers

European Parliament (2020): On automated decision-making processes: ensuring consumer protection and free movement of goods and services https://www.europarl.europa.eu/doceo/document/B-9-2020-0094_EN.pdf

European Data Protection Supervisor (2019): Facial recognition: A solution in search of a problem? https://edps.europa.eu/press-publications/press-news/blog/facial-recognition-solution-search-problem_de

ETIAS (2020): European Travel Information and Authorisation System (ETIAS) https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders/etias_en

ETIAS (2019): European Travel Information and Authorisation System (ETIAS) https://www.eulisa.europa.eu/Publications/Information%20Material/Leaflet%20ETIAS.pdf

Horizon2020 (2019): robusT Risk basEd Screening and alert System for PASSengers and luggage https://cordis.europa.eu/project/id/787120/reporting

High Court of Justice (2019): Bridges v. the South Wales Police https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf

High-Level Expert Group on Artificial Intelligence (2020): Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

Hunton Andrew Kurth (2020): UK Court of Appeal Finds Automated Facial Recognition Technology Unlawful in Bridges v South Wales Police https://www.huntonprivacyblog.com/2020/08/12/uk-court-of-appeal-finds-automated-facial-recognition-technology-unlawful-in-bridges-v-south-wales-police/

Kayalki, Laura (2019): French privacy watchdog says facial recognition trial in high schools is illegal https://www.politico.eu/article/french-privacy-watchdog-says-facial-recognition-trial-in-high-schools-is-illegal-privacy/

Kayser-Bril, Nicolas (2020): EU Commission publishes white paper on AI regulation 20 days before schedule, forgets regulation https://algorithmwatch.org/en/story/ai-white-paper/

Leyen, Ursula von der (2019): A Union that strives for more - My agenda for Europe https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

Leyen, Ursula von der (2020): Paving the road to a technologically sovereign Europe https://delano.lu/d/detail/news/paving-road-technologically-sovereign-europe/209497

Leyen, Ursula von der (2020): Shaping Europe's digital future https://twitter.com/eu_commission/status/1230216379002970112?s=11

Leyen, Ursula von der (2019): Opening Statement in the European Parliament Plenary Session by Ursula von der Leyen, Candidate for President of the European Commission https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_19_4230

Nygård, (2019): The New Information Architecture as a Driver for Efficiency and Effectiveness in Internal Security https://www.eulisa.europa.eu/Publications/Reports/eu-LISA%20Annual%20Conference%20Report%202019.pdf

Sabbagh, Dan (2020): This article is more than 1 month old South Wales police lose landmark facial recognition case https://www.theguardian.com/technology/2020/aug/11/south-wales-police-lose-landmark-facial-recognition-case

South Wales Police(2020): Automated Facial Recognition https://afr.south-wales.police.uk/

Valero, Jorge (2020): Vestager: Facial recognition tech breaches EU data protection rules https://www.euractiv.com/section/digital/news/vestager-facial-recognition-tech-breaches-eu-data-protection-rules/

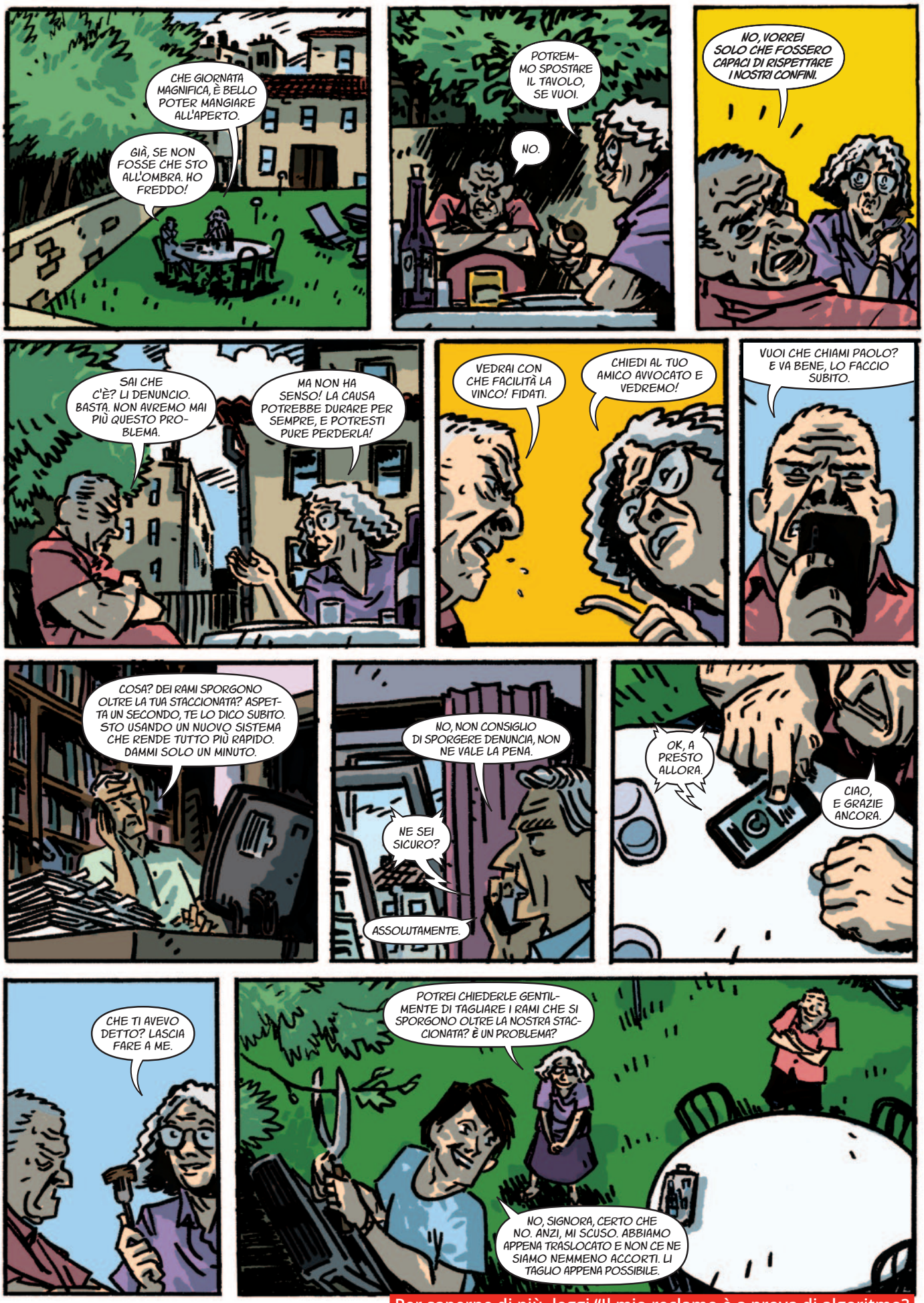
ITALIA

INCHIESTA

PAGINA 37

RICERCA

PAGINA 42



Per saperne di più, leggi "Il mio reclamo è a prova di algoritmo? Giustizia predittiva a Genova (e oltre)" nella sezione dedicata alla "Ricerca".

In Italia c'è voglia di riconoscimento facciale negli stadi

Appena prima dello scoppio della pandemia di COVID-19, il governo e le massime autorità sportive del paese stavano pianificando un'adozione massiva di tecnologie di riconoscimento facciale e sonoro in tutti gli stadi italiani di calcio. Il motivo? Combattere il razzismo.

Di Fabio Chiusi

All'inizio del 2020, proprio mentre il mondo intero si scopriva una buona volta costretto a far fronte alle crescenti evidenze scientifiche sugli esiti discriminatori e razzisti delle tecnologie di riconoscimento facciale, l'Italia ha ventilato l'ipotesi di adottarle in modo capillare negli stadi di calcio, come strumento da impiegare nella lotta contro il razzismo.

La questione aveva trovato ampio spazio sui media mainstream nei mesi precedenti, soprattutto a causa dei ripetuti cori razzisti espressi dai tifosi di Verona e Lazio, tra gli altri, contro la superstar del Brescia, Mario Balotelli, che è di colore.

In febbraio, appena prima che la pandemia di COVID-19 travolgesse l'Italia, il ministro per le Politiche giovanili e lo Sport, Vincenzo Spadafora, ha deciso di averne avuto abbastanza, e pensato che fosse giunto il momento di ricorrere al soccorso della tecnologia. Di conseguenza, Spadafora ha annunciato che il governo italiano sarebbe stato in procinto di sperimentare, e poi adottare, "nuove misure tecnologiche" da mettere a disposizione delle forze di polizia negli stadi di calcio di tutto il paese. E non solo per identificare i soggetti a cui è impedito l'accesso a competizioni sportive: il ministro intendeva impiegarle per individuare i tifosi razzisti, e farlo in tempo reale.

Spadafora ha poi specificato di quali "nuove misure tecnologiche" stesse parlando, al Tg3: riconoscimento facciale e radar sonori. Ma altri lo avevano già fatto per lui. Luigi De Siervo, amministratore delegato della Serie A, aveva parlato già il 16 gennaio del "lavoro silenzioso" condotto dalle istituzioni: "Noi dobbiamo prendere uno per uno le persone che stanno rovinando questo sport meraviglioso. Con il riconoscimento facciale è possibile". Secondo il quotidiano la Repubblica, De Siervo si è perfino recato in Israele per studiarne le più recenti applicazioni.

NOI DOBBIAMO PRENDERE UNO PER UNO LE PERSONE CHE STANNO ROVINANDO QUESTO SPORT MERAVIGLIOSO. CON IL RICONOSCIMENTO FACCIALE È POSSIBILE

Il presidente della Federazione Italiana Giuoco Calcio (FIGC), Gabriele Gravina, aveva a sua volta espresso la medesima idea: "il ricorso alla tecnologia porterà ad un riconoscimento facciale ad alta definizione degli individui che si renderanno colpevoli di questi atti", aveva detto in quello stesso giorno di gennaio. Per Gravina, l'idea avrebbe visto la luce a breve: "In via sperimentale sarà avviata in tempi rapidi, poi sensibilizzeremo tutte le parti in causa, tra società e leghe, nell'utilizzo di questa tecnologia".

/ La polizia vuole ascoltare le conversazioni sugli spalti

Ma non si tratta solo di riconoscimento facciale, come detto: il piano includeva anche l'adozione di sistemi di sorveglianza sonora, i cosiddetti "radar sonori". Già nell'ottobre 2019, Gravina aveva ipotizzato l'impiego dei "radar passivi" attualmente usati dalle unità anti-terrorismo, per captare "con microfoni direzionali la provenienza di un rumore", incluse le conversazioni private tra tifosi sugli spalti — una possibilità che lo stesso Gravina ha definito problematica in termini di privacy. Tuttavia, ciò avrebbe significato anche la possibilità di intercettare conversazioni razziste. I colpevoli sarebbero poi stati identificati tramite riconoscimento facciale.

SOLO IN CASO DI ADOZIONE DELLE TECNOLOGIE ADEGUATE, UN CLUB SAREBBE STATO PROTETTO DA RESPONSABILITÀ DI LEGGE PER GLI ATTI COMMESSI DAI TIFOSI.

L'adozione di queste tecnologie tuttavia avrebbe avuto anche delle conseguenze di ordine legale, secondo Gravina. E infatti, la Federcalcio stava contestualmente pianificando di adottare un nuovo regime di "responsabilità oggettiva" per i club calcistici: solo in caso di adozione delle tecnologie adeguate, questa l'idea, un club sarebbe stato protetto da responsabilità di legge per gli atti commessi dai tifosi. A questo modo, secondo la Federcalcio, le società calcistiche sarebbero state anche liberate dalle estorsioni degli "ultras", che usano spesso il razzismo come arma di ricatto. Per esempio, quando a settembre 2019 la Juventus ha smesso di prestarsi alle loro presunte estorsioni — per ottenere abbonamenti stagionali, bevande gratuite ai bar dello stadio, e perfino inviti a feste esclusive —, gli ultras hanno cominciato a levare canti razzisti come rappresaglia,

ben consci che si sarebbero tradotti in sanzioni per la società. Risultato? Dodici ultras sono stati arrestati.

Anche se alcuni media hanno scritto che perfino diversi membri dello staff del ministro degli Interni nutrissero seri dubbi circa l'efficacia di tali tecnologie in contesti di vita reale, altri le hanno lodate a prescindere, definendole "la nuova frontiera della lotta al razzismo". De Siervo si è perfino paragonato all'ex premier britannica, Margaret Thatcher, dicendo: "Vogliamo riuscire a fare in due anni quello che la Thatcher ha fatto in dieci".

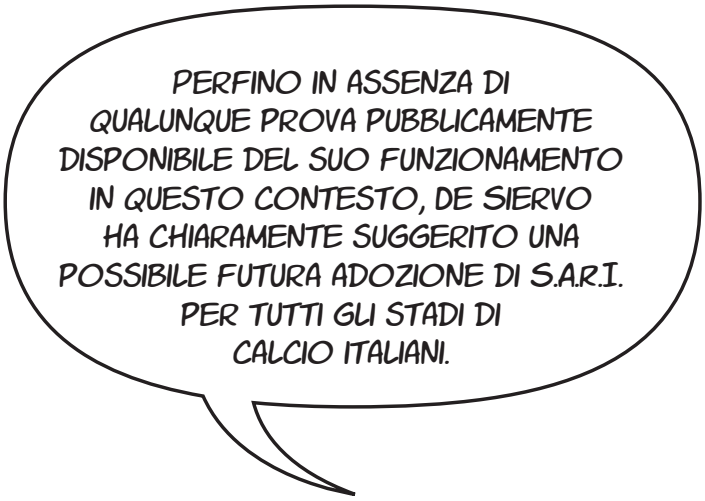
Poi la pandemia si è abbattuta sull'Italia, ed è cambiato tutto.

/ Il governo fornisce una vaga conferma dei suoi piani, gli altri tacciono

Mentre l'epidemia di COVID-19 andava spandendosi a macchia d'olio nel Nord Italia, il governo, il 10 marzo 2020, ha deciso di sospendere il campionato di Serie A. Gli stadi sarebbero inizialmente rimasti chiusi al pubblico fino a giugno, ritardando così a data da destinarsi la pianificata adozione delle "nuove misure tecnologiche" ipotizzate.

E tuttavia, in uno scambio via mail a maggio 2020, un portavoce del ministro per le Politiche giovanili e lo Sport ha "certamente" — anche se senza fornire ulteriori dettagli — confermato di stare "cercando di implementare una serie di strumenti di prevenzione dei fenomeni di razzismo negli stadi e che sicuramente siamo a lavoro con la FIGC per collaborare sulla questione". Nessun aggiornamento è stato ritenuto necessario a seguito di un ulteriore scambio via mail a settembre, e nessuna ulteriore risposta è giunta alle precise domande poste da AlgorithmWatch riguardo ai piani del governo.

Nemmeno Lega Calcio Serie A e Federcalcio hanno risposto alle nostre domande. E tuttavia, ciò non ha impedito ad alcuni club di condurre sperimentazioni di tecnologie di riconoscimento facciale nei loro stadi. Grazie ad articoli giornalistici siamo venuti a conoscenza di progetti pilota sia a Udine che a Napoli, dove 190 videocamere dotate di riconoscimento facciale sono attive fin da settembre 2019. Un mese dopo, la videosorveglianza dello stadio San Paolo era già stata cruciale nell'identificazione di 32 tifosi a cui è stata comminata una multa di 500 euro ciascuno — 166 se corrisposti entro cinque giorni dalla ricezione della notifica — per violazioni delle regole di sicurezza dello stadio partenopeo.



/ Altro lavoro per S.A.R.I.

Ma è la sperimentazione tenutasi a Udine ad avere introdotto lo sviluppo più rilevante. Come descritto nel dettaglio nella prima edizione di questo rapporto, un "Sistema Automatico di Riconoscimento Immagini", cioè il "S.A.R.I.", era già stato utilizzato dalla polizia italiana per arrestare trasgressori della legge, sia confrontando l'immagine facciale di un sospetto con quelle incluse nei database gestiti dalle forze dell'ordine (funzione "ENTERPRISE"), sia consentendo il riconoscimento facciale da video in tempo reale (funzione "REAL TIME").

Concepito nel 2014, e adottato in diversi contesti (dal porto di Bari a un casinò a Las Vegas), il sistema è stato testato nel giugno 2019 allo Stadio Friuli di Udine, prima della finale degli Europei Under 21. L'obiettivo era metterlo alla prova in funzione di monitoraggio agli ingressi, così da lasciare fuori dallo stadio i soggetti che non hanno diritto di entrarvi, per esempio per via di provvedimenti restrittivi (DASPO).

AlgorithmWatch ha chiesto ripetutamente sia all'Udinese FC che a Reco 3.26, i creatori — di base in Puglia — di S.A.R.I., di poter accedere ai risultati della sperimentazione, ma la società calcistica ha delegato ogni risposta agli sviluppatori, e gli sviluppatori non hanno mai risposto ad alcuna domanda.

Ciononostante, e perfino in assenza di qualunque prova pubblicamente disponibile del suo funzionamento in questo contesto, De Siervo ha chiaramente suggerito una possibile futura adozione di S.A.R.I. per tutti gli stadi di calcio italiani, affermando esplicitamente che il sistema da lui individuato per realizzare questa ristrutturazione tecnologica della sicurezza negli stadi sarebbe stato proprio "lo stesso che ha la polizia".

La completa mancanza di un qualunque dibattito pubblico in materia non ha impedito ai media mainstream di dipingere S.A.R.I. come una misura di sicurezza predefinita, da adottare a prescindere, per i piani sia dell'Inter che del Milan per il nuovo stadio che andrà a rimpiazzare lo storico San Siro nel capoluogo lombardo. Qui l'idea è di non limitarsi ai controlli per gli individui a cui è bandito l'ingresso, ma di estenderli anche a quelli per individuare "possibili terroristi", ha scritto il Corriere della Sera Milano, promettendo di consentire il riconoscimento di un volto "entro 60 metri" di distanza.

Ma, di nuovo, c'è di più. Secondo lo stesso articolo, il piano condiviso con le autorità cittadine da Inter e Milan include anche "geolocalizzazione e audiorilevatori", insieme a "software in grado di segnalare comportamenti anormali", tra cui il "vagabondaggio" (loitering) e "la comparsa improvvisa di oggetti in un determinato spazio".

L'ex Garante Privacy invita alla prudenza

Razzismo. Terrorismo. Vagabondaggio. Il riconoscimento facciale negli stadi ha molte facce — e perfino più impieghi. Videocamere dotate di riconoscimento facciale sono state installate, per esempio, nel nuovo stadio dell'Atalanta, il "Gewiss Stadium" di Bergamo. Qui, infatti, la ristrutturata "Curva Pisani" dispone ora di 7 fornelli e 40 videocamere di sorveglianza, alcune delle quali dotate di riconoscimento facciale. Già nel 2016, del resto, il Garante italiano per la privacy ne aveva autorizzato l'uso anche allo Stadio Olimpico di Roma.

Ma le cose sono molto cambiate, rispetto ad allora. Negli ultimi anni le evidenze scientifiche sui risultati discriminatori prodotti dalle tecnologie di riconoscimento facciale sono andate moltiplicandosi. E quelle prove dovrebbero essere tenute in considerazione, argomenta l'ex presidente dell'Autorità Garante per la protezione dei dati personali italiana, Antonello Soro, in una intervista scritta con AlgorithmWatch, condotta quando era ancora in carica.

Nell'intervista, Soro riconosce che "il contesto, assolutamente peculiare, degli stadi, è tra quelli nei quali si sono manifestate con maggiore evidenza spinte all'utilizzo esteso della biometria". Di conseguenza, la previsione di adeguate tutele è ancora più importante. E, dato che, rispetto al 2016, i "peculiari rischi" derivanti dal controllo biometrico sono nel frattempo divenuti evidenti — "in particolare se funzionale a processi decisionali di tipo algoritmico" —, Soro chiede "ulteriore cautela" prima di ricorrerne all'utilizzo. Più nel dettaglio, ciò significa che i sistemi di riconoscimento facciale e di controllo biometrico devono essere armonizzati alla normativa esistente, e che è richiesta inoltre una valutazione del loro impatto sulla privacy dei cittadini. È di questo che il Garante ha discusso anche coi vertici della FIGC, afferma Soro.

E tuttavia, la funzione "REAL TIME" di S.A.R.I., quella che risulterebbe rilevante per i piani antirazzismo del ministro Spadafora, non è stata oggetto di "interlocuzioni preliminari con il Ministero dell'interno" durante il suo mandato, scrive Soro. La proposta non risulta così essersi "tradotta in un progetto definitivo" — principalmente per l'assenza di una "previsione normativa adeguata", che fosse dunque capace di declinare "nel dettaglio limiti e garanzie" circa il trattamento dei dati biometrici ottenuti, e di farlo con le cautele che vanno riservate a un caso "indubbiamente significativo per dimensioni e tipologia delle operazioni realizzate" come quello in esame.

L'attuale, nuova presidenza dell'Authority, insediatasi a luglio 2020, non ha ritenuto di dover fornire ad AlgorithmWatch alcun aggiornamento in materia.

/ La pandemia come pretesto per nuovi usi del riconoscimento facciale

In assenza di chiare regale di trasparenza, i test e le implementazioni di tecnologie di riconoscimento facciale hanno cominciato a moltiplicarsi, al riparo dallo scrutinio del pubblico. Poi, però, è arrivata la pandemia di COVID-19, e di conseguenza le gare sono state sospese, svuotando gli spazi pubblici.

A quel punto, si è posta una questione elementare: quando gli stadi riapriranno, i tifosi dovranno indossare una mascherina, per entrarvi. Ma le mascherine riducono sensibilmente l'accuratezza delle tecnologie di riconoscimento facciale, come documentato da un recente studio del National Institute of Standards and Techno-

NEGLI ULTIMI ANNI LE EVIDENZE SCIENTIFICHE SUI RISULTATI DISCRIMINATORI PRODOTTI DALLE TECNOLOGIE DI RICONOSCIMENTO FACCIALE SONO ANDATE MOLTIPLICANDOSI.

logy (NIST) statunitense. Che fare? Possono le tecnologie di riconoscimento facciale venire migliorate al punto di riconoscere anche gli individui che indossano una mascherina?

Diverse aziende affermano di aver risolto il problema, ma le evidenze scientifiche attualmente disponibili suggeriscono altrimenti. È forse per questo che le sperimentazioni attuali (così come il materiale pubblicitario delle aziende produttrici) si concentrano principalmente sul semplice riconoscimento facciale, senza vantare insieme l'affermazione, ambiziosa, di essere in grado di riconoscere anche gli individui che indossano una mascherina. Chi lo fa, poi, vi aggiunge di norma tecnologie per la misurazione della temperatura corporea (termoscanner) e altri strumenti tecnologici impiegati per tenere sotto controllo lo stato della salute pubblica durante una pandemia.

Un esempio viene da un progetto pilota tenutosi a Torino, in cui, in occasione della partita tra Torino e Hellas Verona, è stato sperimentato un "varco antivirus automatizzato" chiamato sistema "Feel Safe". "Feel Safe" è stato progettato per velocizzare l'ingresso dei tifosi in sicurezza, ed è prodotto dall'azienda milanese Worldwide Exhibition Systems (WES). Non sono risultate reperibili informazioni circa le tipologie di dati biometrici prese in considerazione dal sistema, nemmeno nel suo materiale aziendale o pubblicitario. Una presentazione di WES dice solamente che, "attraverso un sistema di controllo biometrico", Feel Safe può "interfacciarsi con il sistema di gestione degli accessi aziendali".

L'Italia, naturalmente, non è la sola a sperimentare tecnologie di riconoscimento facciale negli stadi. E se Danimarca, Belgio e Olanda le hanno impiegate per i controlli di sicurezza — con risultati altalenanti e non senza ricorsi —, sono state anche ipotizzate come strumento "per fornire esperienze personalizzate ai tifosi" nello stadio dei Dodger, a Los Angeles.

Nessun altro paese oltre all'Italia, tuttavia, ha cercato di applicare sistemi di riconoscimento facciale e sonoro all'identificazione degli autori di insulti razzisti.

Che questa idea finisca prima o poi per venire realmente implementata è tutto da vedere. Noi, in ogni caso, saremo qui, a controllare.

Ricerca

Di Fabio Chiusi

Il contesto dell'ADM in Italia

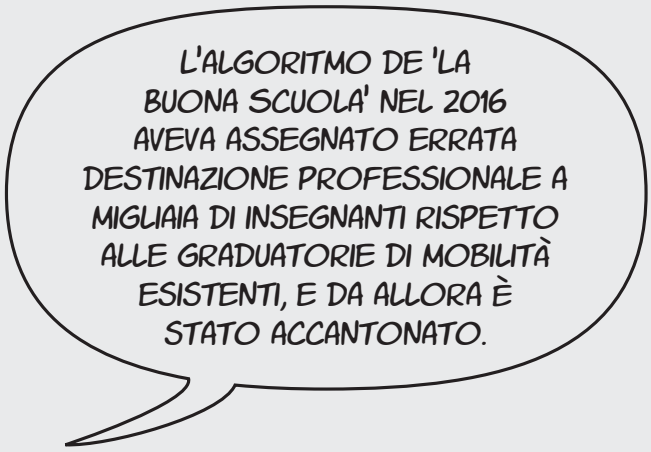
/ Più dibattito, ma sull'intelligenza artificiale

Sebbene non ancora riconosciuti come tali, i sistemi per automatizzare processi decisionali ("automated decision-making", o "ADM") sono sempre più rilevanti nei dibattiti pubblici e sulle scelte politiche in Italia. Mancano norme specifiche, e — come notato nella prima edizione di questo rapporto — non c'è una definizione formale di "ADM" né per gli usi governativi né per quelli privati. Eppure, due sentenze pronunciate nel corso del 2019 prescrivono ora limiti più chiari circa l'utilizzo di sistemi di ADM nel sistema educativo. Il riferimento è all'algoritmo de 'La Buona Scuola', che nel 2016 aveva assegnato errata destinazione professionale a migliaia di insegnanti rispetto alle graduatorie di mobilità esistenti, e che da allora è stato accantonato.

I sistemi di ADM sono al cuore del dibattito sorto intorno all'uso dei dati per combattere la pandemia di COVID-19, attraverso una applicazione per smartphone che ha ricevuto considerevole copertura mediatica, e anche in relazione all'uso di analisi predittive di Big Data per verificare la coerenza tra dichiarazioni al fisco e reale posizione fiscale tramite gli Indici Sintetici di Affidabilità (ISA), implementati dal governo in carica.

Anche la polizia predittiva ha lasciato il segno in Italia, e i risultati dettagliati in questo rapporto mostrano chiaramente come soluzioni quali XLAW siano già diffusamente testate, adottate e implementate. Anche progetti di giustizia e sanità predittiva sono in corso di sperimentazione, perlopiù ancora a livello locale.

Tuttavia, per la maggior parte degli italiani, dibattiti come questi riguardano l'intelligenza artificiale (AI), non i sistemi per automatizzare processi decisionali (ADM). Le implicazioni sociali e politiche dell'AI informano ancora, di conseguenza, buona parte della retorica mainstream circa le sfide e le opportunità derivanti dall'automazione, con i media tradizionali troppo spesso intenti a indugiare su panici morali e articoli acchiappa-click (*click-bait*).



Allo stesso tempo, la dimensione pubblica dei dibattiti sulle politiche tecnologiche riguardanti equità, discriminazioni e manipolazioni algoritmiche procede altrettanto spesso senza riguardo alcuno per fatti ed evidenze scientifiche, nonostante in materia vadano accumulandosi sia nella letteratura accademica che tramite inchieste giornalistiche.

/ Il riconoscimento facciale è improvvisamente ovunque

Ciò potrebbe contribuire a spiegare perché la crescente adozione di soluzioni di ADM invasive — quali sono le tecnologie di riconoscimento facciale — non sia ancora debitamente apprezzata dalla maggioranza degli italiani, nonostante un rapporto di Comparitech (Forbes Italia 2019) abbia, nell'ottobre del 2019, redarguito l'Italia per "l'esteso uso di tecnologie biometriche, incluso il riconoscimento facciale negli aeroporti", e aggiungendo come ciò stesse già "causando preoccupazioni tra i cittadini".

Mentre il riconoscimento facciale diventava noto alle masse tramite applicazioni per smartphone di successo, come

FaceApp, si sono anche moltiplicate le sperimentazioni per farvi ricorso in luoghi pubblici e istituzionali, negli stadi, negli ospedali e nelle scuole. Inoltre, come se non bastasse, ne vengono pianificate di ulteriori, pur in assenza di un vero e proprio dibattito democratico.

A tutto questo non si è opposta alcuna campagna di attivismo realmente significativa, consentendo così che questi esperimenti procedessero perlopiù in assenza di contraddittorio e, perfino quando contestati da giornalisti ed esperti, con una mancanza sconcertante di trasparenza. Ciò è massimamente evidente nel caso del sistema — sempre in evoluzione — di riconoscimento facciale S.A.R.I., in uso dalle forze dell'ordine: il suo misterioso funzionamento non è ancora stato esaustivamente spiegato al pubblico.

Questa mancanza di dettagliato scrutinio ha inoltre portato a una crescente tendenza a dipingere gli algoritmi in maniera "distopica", e a (ingiustamente) identificarli come la causa diretta di molti dei difetti attuali della società italiana. È agli algoritmi, infatti, che vengono regolarmente imputate polarizzazione, mancanza di fiducia nelle istituzioni e nei media, "fake news" — automatizzate o meno — e perfino la crisi della democrazia. Inoltre, i crescenti segnali di razzismo, odio (*hate speech*), nazionalismo di estrema destra, e intolleranza sono stati a loro volta sostanzialmente attribuiti agli algoritmi di personalizzazione e microtargeting delle piattaforme digitali. E di nuovo, tutto ciò è avvenuto perlopiù in assenza di prove certificate dalla comunità scientifica (*peer-reviewed*) o di solide inchieste circa i reali effetti delle campagne sui social media in termini di persuasione.

/ A cosa pensano gli italiani quando pensano all'ADM

Ma a cosa pensano davvero gli italiani, quando pensano all'ADM? La migliore approssimazione attualmente disponibile è cercare di comprendere cosa pensino dell'AI in contesti di decision-making. Nel 2019, due sondaggi hanno cercato di indagare il ruolo dell'intelligenza artificiale nelle vite quotidiane degli italiani sotto diversi aspetti, e alcuni dei risultati ottenuti sono rilevanti per il dibattito sull'ADM.

Un primo sondaggio di AIDP-Lablaw, condotto da BVA DOXA e pubblicato a novembre 2019 (BVA DOXA 2019), sostiene che gli italiani sono in generale speranzosi che l'AI porti a soluzioni e scoperte altrimenti inconcepibili (94%), e che credono in larga misura che l'AI arrechi benefici in termini di benessere e qualità della vita (87%). Inoltre, l'89% ritiene che i robot non rimpiazzeranno mai del tutto l'intervento umano — e, pur se in massima parte riferito al mercato del lavoro, ciò può essere considerato di rilievo anche in termini di ADM.

In ogni caso, gli italiani interpellati temono ancora che l'AI finirà per avere un ruolo dominante nelle loro vite (il 50% crede ci sia il rischio che le macchine finiranno per sottomettere l'uomo), a un'ampia maggioranza (il 70%) ritiene che "molti posti di lavoro" saranno persi in conseguenza a un incrementato ricorso all'automazione. Un risultato cruciale è il sorprendente consenso in tema di regolamentazione dell'AI: il 92% si è detto d'accordo con l'idea che delle leggi siano necessarie.

Un secondo sondaggio, sempre pubblicato a novembre 2019, ha poi cercato di affrontare una questione più precisa: gli italiani sono preoccupati dai "robot killer" — e, di conseguenza, da ciò che potremmo definire "l'ADM nelle guerre"? Secondo il sondaggio YouGove contenuto nella campagna internazionale per "fermare i robot killer", gestita in Italia da Rete Disarmo (Rete Disarmo 2019), il 75% degli italiani è in favore di un divieto internazionale allo sviluppo di sistemi completamente autonomi che siano attrezzati per uccidere esseri umani, e si spinge fino a proporre che l'Italia si ponga come paese leader nell'ottica di creare un sistema di norme internazionali in materia.

Un precedente sondaggio IRIAD Archivio Disarmo (marzo 2019) aveva mostrato risultati simili. L'idea di un divieto è inoltre supportata in modo trasversale, nell'intero spettro partitico e ideologico.

Un altro sondaggio meno recente, di monster.it, sull'automazione nei processi di assunzione è rilevante per meglio comprendere l'atteggiamento degli italiani nei confronti dell'ADM. Pubblicato a dicembre 2018, il sondaggio concludeva (Sole 24 Ore 2018) che quasi metà del campione (il 45%) credesse fermamente che "assolutamente no, un robot non potrà mai valutare una persona meglio di un essere umano". Tuttavia, quasi un quarto dei rispondenti riteneva anche che gli algoritmi avrebbero potuto — e sarebbero stati — di aiuto nel fornire assistenza alla valutazione dei candidati. Il 6% si è detto d'accordo con l'affermazione secondo cui gli algoritmi diventeranno talmente affidabili e precisi da poter sostituire gli esseri umani.

Un catalogo di esempi di ADM

/ Il tuo volto è la carta d'imbarco: riconoscimento facciale negli aeroporti

Tecnologie di riconoscimento facciale stanno venendo adottate, o stanno per essere adottate, in entrambi gli aeroporti di Milano, Linate e Malpensa, e a Roma Fiumicino — tre dei più importanti snodi aeroportuali italiani, sia per i voli nazionali che per quelli internazionali. L'idea è che tramite queste tecnologie si possa consentire ai viaggiatori di avere finalmente un'esperienza priva di qualunque lungaggine o intoppo ("*frictionless journey*"), rimpiazzando identificazione e carta d'imbarco con il volto del passeggero. Le autorità aeroportuali sostengono che queste tecnologie renderanno i controlli più rapidi, incrementandone al contempo il livello di sicurezza.

Il 12 febbraio 2020, un progetto pilota basato su tecnologie di riconoscimento facciale è stato avviato all'aeroporto di Linate, con l'obiettivo di durare fino a fine anno (Corriere Milano 2019; Corriere della Sera 2020). Chiamato "Face Boarding", il progetto coinvolge, almeno inizialmente, i soli voli Alitalia da Fiumicino, e solo per passeggeri abituali (*frequent flyers*) e in classe "Comfort" (cioè "business"). L'inizio di una sperimentazione è previsto anche a Malpensa entro la fine del 2020 (Mobilità 2019).

All'aeroporto di Roma Fiumicino una sperimentazione è cominciata nel novembre 2019, a partire da una sola compagnia, KLM, e una sola destinazione, Amsterdam (TgCom24 2019). Tuttavia, il management dell'aeroporto ha affermato che il progetto sarà presto ampliato, e che finirà per applicarsi a tutte le destinazioni e a tutti i voli in partenza e in arrivo all'aeroporto internazionale. Gli sviluppatori del sistema, realizzato attraverso una collaborazione tra ADR (Aeroporti di Roma) e il fornitore di soluzioni per l'identità elettronica Vision-Box, affermano di raccogliere i dati biometrici dei viaggiatori solo temporaneamente, e che tutti i dati saranno distrutti "entro un'ora dalla partenza": "niente rimarrà sui server di ADR", scrivono sul sito dell'aeroporto.

Il restyling tecnologico a base di biometria di entrambi gli hub milanesi di Malpensa e Linate costerà 21 milioni di euro, ha scritto il canale all-news TgCom24, e includerà "sei chioschi per l'enrollment (l'iscrizione per associare i dati biometrici a passaporto o carta d'identità elettronici), 25 boarding gate elettronici, 10 gate pre-security e 7 face spot security".

I gate per controlli automatici a base di riconoscimento facciale (chiamati "Automated Border Control Gates", o "ABCGates") erano già stati installati all'aeroporto di Napoli nel 2016, gestendo il passaggio di circa duemila passeggeri al giorno (Repubblica Napoli 2016). "La nuova tecnologia ha contribuito a migliorare l'esperienza dei passeggeri all'Aeroporto di Napoli", ha affermato Alessandro Fidato, Direttore Infrastructure & Operations dell'aeroporto, "velocizzando il transito ai controlli e assicurando, nello stesso tempo, il rispetto dei più alti standard di sicurezza alle frontiere italiane". La soluzione informatica è fornita da Sita, descritta nei media come "leader mondiale" nella tecnologia per la sicurezza aeroportuale. L'azienda afferma di stare già rifornendo all'incirca trenta governi in tutto il mondo.

Secondo indiscrezioni giornalistiche, ulteriori sperimentazioni saranno inevitabili, specialmente a seguito della pandemia di COVID-19 (Il Messaggero 2020).

NESSUN AUDITING INDIPENDENTE O ALTRO RISULTATO ANALITICO RIGUARDANTE QUESTI SISTEMI DI RICONOSCIMENTO FACCIALE È PUBBLICAMENTE DISPONIBILE, E NESSUNO DI ESSI È STATO ADEGUATAMENTE DISCUSSO DALL'OPINIONE PUBBLICA.

Nessun auditing indipendente o altro risultato analitico riguardante questi sistemi di riconoscimento facciale è pubblicamente disponibile, e nessuno di essi è stato adeguatamente discusso dall'opinione pubblica .

/ Questi sistemi automatici predicono la tua salute

Gli esperimenti di sistemi di ADM per la "salute predittiva" (*predictive health*) non sono inediti, in Italia. A quelli descritti nella prima edizione di questo rapporto, 'RiskER' e 'Abbiamo i numeri giusti', si è infatti aggiunto un altro interessante esperimento nel corso del 2019. Nella città di Vimercate, in Lombardia, il ramo locale dell'Azienda Socio Sanitaria Territoriale (ASST) ha adottato soluzioni cloud a codice aperto (*open source*), sviluppate da Almaviva, che usano algoritmi per predire l'insorgenza di patologie croniche e complicazioni post-operatorie (Giornale Monza 2019).

Dopo avere trascorso un decennio a digitalizzare cartelle sanitarie, l'ospedale di Vimercate mira ora a farne la base per una più efficiente gestione dei pazienti, anche attraverso la personalizzazione delle cure sulla base di indicazioni fornite da analisi di machine learning, per ridurre così i costi e ottimizzare la logistica, migliorando al contempo le proprie soluzioni di "medicina di precisione".

"È la prima esperienza italiana di utilizzo dell'AI in modo strutturato in ambito ospedaliero", dice l'Head of AI Solutions e IoT Platform di Almaviva, Antonio Cerqua, che ha anche affermato che l'obiettivo è di "scalare" la piattaforma di Internet delle cose (*Internet of Things*, o "IoT") della propria compagnia "in un network di livello nazionale" tra ospedali interessati. (Il Giorno 2019)

Esperti del colosso della consulenza Deloitte hanno a loro volta visitato l'ospedale di Vimercate proprio per interessarsi al progetto, hanno scritto il Giorno e il Giornale di Monza.

"L'AI non deve sostituirsi ai medici", dice Giovanni Delgrossi, direttore dei Sistemi informativi dell'Asst della cittadina lombarda, "ma essere un supporto, uno strumento che aiuta a prendere le decisioni migliori". (Sole 24 Ore 2019b)

/ Il mio reclamo è a prova di algoritmo? Giustizia predittiva a Genova (e oltre)

Nel 2019, un primo esperimento di ADM nel sistema giudiziario italiano è stato sviluppato dal LIDER Lab della Scuola Superiore Sant'Anna, a Pisa, in collaborazione con EMbeDS, KDD Lab e il Tribunale di Genova (Gonews 2019).

Chiamato 'Predictive Jurisprudence' ("giurisprudenza predittiva"), il progetto consente ai ricercatori di avere accesso all'intero corpus delle sentenze emesse dal sistema giudiziario nel capoluogo ligure, e analizzarlo attraverso tecniche di machine learning. Lo scopo è estrarre informazioni rilevanti per ulteriori analisi, a partire dall'identificazione di tendenze comuni nella giurisprudenza riguardo a uno specifico caso in esame (per esempio, sentenze passate su certe tipologie di crimini o altri casi giudiziari analoghi).

A questo modo qualunque giudice in carne e ossa disporrebbe di un punto di riferimento per meglio valutare il caso in esame, e insieme di uno strumento con cui verificare facilmente la coerenza tra la propria posizione e quella emersa in sentenze precedenti in circostanze analoghe.

Il presidente del Tribunale di Genova, Enrico Ravera, sostiene che lo scopo del progetto non sia la mera estrapolazione di "statistiche dall'analisi casistica", quanto piuttosto applicare "tecniche di intelligenza artificiale alla giurisprudenza".

Più di preciso, 'Predictive Jurisprudence' è descritto dal team di ricercatori che conduce il progetto, guidato dal prof. Giovanni Comandé, come un "progetto a più livelli che si sviluppa su cinque livelli autonomi e interconnessi".

Scrivono Comandé e la collega Prof.ssa Denise Amram in una email di risposta alle domande di AlgorithmWatch: "Nella fase di startup, il progetto mira ad analizzare i giudizi, e i relativi documenti processuali, secondo i criteri e le metodologie sviluppate nell'Observatory on Personal Injuries, applicabili ad aree di contenzioso diverse da quelle riguardanti danni non-monetari (livello 1). Lo stesso materiale viene anche analizzato tramite tecniche di machine learning per sviluppare sia strumenti per l'annotazione e l'estrazione automatica di informazioni da testi di legge (livello 2) sia algoritmi per l'analisi e la predizione (cosiddetto "livello 3 intelligenza artificiale").

In particolare, scrivono i ricercatori, "l'architettura del database costruito per ospitare i dati acquisiti dalle corti sarà concepita e addestrata per sviluppare algoritmi che identifichino automaticamente tendenze secondo criteri noti all'interprete, e che allo stesso tempo mettano in evidenza nuovi trend sulla base di possibili bias e tendenze scoperte dall'algoritmo stesso".

Particolarmente degno di nota in termini di ADM è il fatto che "gli algoritmi mirino a ricreare e mimare i ragionamenti legali alla base dell'adozione di certe misure nelle sentenze, rendendo prevedibili ulteriori decisioni future sulla stessa materia". L'automazione viene, di conseguenza, applicata al riassunto e la spiegazione di idee: "Questi strumenti dovrebbero anche aiutarci a spiegare la ratio dietro a ogni decisione, proprio mentre sperimentiamo lo sviluppo di strumenti adatti a spiegare i criteri definiti dall'AI (livello 4)".

Infine, nell'ultima fase, "gli sforzi intrapresi e i risultati ottenuti nei differenti livelli di ricerca e sviluppo saranno ricondotti al tentativo di strutturare un'analisi dell'argomento legale a un livello di astrazione e sistematicità tali da contribuire alla semplificazione di tutti i compiti (livello 5)".

Tra i casi di studio attuali, è inclusa una prima sperimentazione in tema di "assegno di mantenimento in caso di divorzio". In questa circostanza, "le domande sono predetermi-

nate dalla legge, ma la loro interpretazione giurisprudenziale evolve di continuo", scrivono i ricercatori. "A questo proposito, alcune riforme attuali propongono di introdurre nuovi criteri, la cui efficacia potrebbe essere discussa alla luce delle nostre analisi".

"Il secondo e il terzo esperimento", scrivono Comandé e Amram, "vengono sviluppati all'interno dell'Observatory on Personal Injury Damage Studies. L'algoritmo può contribuire all'individuazione di criteri per dispensare compensazioni per perdite non monetarie che vadano oltre le interpretazioni correnti, e tentino di uniformare il trattamento di questa forma altamente soggettiva di danni. All'interno di questa fondamentale analisi, l'algoritmo potrebbe essere meglio addestrato a spiegare le perdite non monetarie nei casi di esaurimento, i cui confini stanno al momento venendo ancora discussi sia dalla prospettiva clinica che da quella legale".

Secondo i ricercatori, "diversi Tribunali stanno si stanno aggiungendo al progetto. Tra di essi, il Tribunale di Pisa e quello di Bologna".

Nel 2018, anche la Corte di Appello e il Tribunale di Brescia hanno avviato un esperimento di "giustizia predittiva", attraverso un progetto pilota il cui obiettivo è ricavare previsioni circa sia la lunghezza di una causa, che i principi che saranno con maggiore probabilità adottati nel valutarla. Il sistema dovrebbe spingersi fino a fornire una stima di probabilità di vedere un reclamo approvato. (BresciaOggi 2018).

La sperimentazione è parte di una collaborazione della durata di due anni tra il sistema giudiziario e l'Università di Brescia, e si compone dei seguenti passi: l'identificazione della questione da analizzare; la creazione di un database che contenga ogni argomento analizzato; la costruzione dei "gruppi di lavoro", tra le risorse a disposizione dell'Università, che si relazioneranno operativamente con il Tribunale; l'estrazione di previsioni da parte dei ricercatori; e, infine, la pubblicazione e diffusione dei risultati di ricerca (AgendaDigitale 2019) — che, mentre scriviamo, non sono ancora disponibili.

Scelte politiche, forme di controllo, e dibattiti sull'ADM

/ Lezioni dal fallimento dell'algoritmo de La Buona Scuola

Alcuni degli sviluppi più importanti verificatisi nel campo dell'ADM nel corso dell'ultimo anno riguardano i modi in cui le corti italiane hanno risposto alle cause intentate dai docenti finiti, loro malgrado, vittima degli errori sistematici compiuti dall'algoritmo de La Buona Scuola. L'algoritmo era stato usato nel 2016 per valutare le richieste di mobilità di 210.000 insegnanti. (Repubblica, 2019) La valutazione e l'applicazione delle procedure di gestione delle domande erano state interamente delegate a un sistema di ADM fallace.

Da allora, l'algoritmo è stato accantonato. E tuttavia, i ricorsi contro le decisioni prese dal sistema di ADM si contano ancora nell'ordine delle migliaia (gli insegnanti colpiti sarebbero 10.000, secondo la Repubblica). E, come dettagliato nella prima edizione di questo rapporto, tali ricorsi hanno una solida base fattuale — dalle pessime scelte di programmazione adottate, alle strategie di offuscamento messe in atto per impedire un qualunque reale auditing del codice sorgente dell'algoritmo de La Buona Scuola.

Il risultato? Docenti spostati dalla Puglia a Milano, invece che all'interno della propria regione, o sottratti alla loro sede naturale a Padova, senza motivo. Il sistema ha perfino costretto due insegnanti con figli autistici a spostarsi dalla Calabria a Prato, nel nord della Toscana.

Nel corso del 2019, due sentenze in particolare hanno precisato come e perché questi errori automatizzati si siano tradotti in scelte e discriminazioni illecite riguardanti le vite professionali e personali di migliaia di individui impiegati nel sistema dell'istruzione pubblica. La prima viene dal Consiglio di Stato (State Council 2019), e la seconda dalla Sezione Terza-Bis del Tribunale Amministrativo (TAR) del Lazio (n. 10964, Settembre 2019). (TAR Lazio 2019)

Per quanto entrambe le istituzioni riconoscano le falle tecniche che hanno funestato la prestazione dell'algoritmo di categorizzazione, esse allo stesso tempo chiariscono che la vera falla sia anche più fondamentale, dato che affligge la

logica stessa che ha portato all'adozione di quell'algoritmo — secondo il TAR, finendo addirittura in contraddizione con la Costituzione italiana e la Convenzione europea dei diritti dell'uomo.

Infatti, anche se il Consiglio di Stato rammenta che l'automazione dei processi decisionali all'interno di un'amministrazione pubblica apporta "indiscutibili vantaggi", quest'ultimo prescrive anche che quei vantaggi possono realizzarsi solo se tali processi non includono alcun "apprezzamento discrezionale" da parte del sistema. In particolare, le sue formule tecniche di programmazione e funzionamento, e le decisioni da esso prodotte, devono essere sempre, in ultima analisi, il risultato di giudizio umano. Gli algoritmi possono anche decidere da soli, ma solo nel caso in cui le decisioni siano in realtà procedure altamente standardizzate e ripetitive.

Anche il TAR del Lazio ha concluso che "le procedure informatiche, finanche ove pervengano al loro maggior grado di precisione e addirittura alla perfezione, non possano mai soppiantare, sostituendola davvero appieno, l'attività cognitiva, acquisitiva e di giudizio che solo un'istruttoria affidata ad un funzionario persona fisica è in grado di svolgere". Cruciale, di nuovo, è il fatto che l'intera procedura sia stata delegata a un "algoritmo impersonale", quando invece l'automazione può unicamente avere un ruolo "strumentale", o "meramente ausiliario" all'interno del procedimento amministrativo. Il TAR precisa inoltre chiaramente che agli algoritmi non può essere "giammai" imposto un ruolo "dominante o surrogatorio dell'attività dell'uomo".

I RICORSI CONTRO LE DECISIONI PRESE DAL SISTEMA DI ADM SI CONTANO ANCORA NELL'ORDINE DELLE MIGLIAIA.

Ciò ha evidenti implicazioni in termini di trasparenza e spiegabilità delle ragioni alla base delle decisioni delle amministrazioni pubbliche, come sottolineato in entrambe le sentenze. E tuttavia, le cornici interpretative sono diverse. Il Consiglio di Stato non esclude la possibilità di decisioni completamente automatizzate e, insieme, trasparenti. A fondamento della sua risposta favorevole ai ricorrenti, infatti, il Consiglio ha addotto unicamente il fatto che l'algoritmo utilizzato da La Buona Scuola fosse talmente compromesso e caotico da risultare "inconoscibile" — sia per i docenti che ne sono rimasti vittima, sia per il Consiglio di Stato stesso. Ciononostante, la sentenza aggiunge immediatamente che, nel caso di piena automazione di una decisione, andrebbe applicata una declinazione rafforzata del concetto di "trasparenza". Una declinazione che implica anche il principio "della piena conoscibilità di una regola espressa in un linguaggio differente da quello giuridico".

Il TAR del Lazio, invece, definisce la possibilità stessa di tale piena automazione una "deleteria prospettiva orwelliana", suggerendo inoltre che trasparenza e spiegabilità ne risulterebbero, di conseguenza, compromesse alla radice.

/ Applicazione degli indici algoritmici di affidabilità fiscale

Introdotti dal governo Gentiloni nel 2017, gli Indici Sintetici di Affidabilità (ISA) sono entrati in vigore nel 2019. Sviluppato da SOSE (Soluzioni per il sistema economico), una partecipata del Ministero dell'Economia e delle Finanze (MEF) e di Banca D'Italia, il sistema consiste in una valutazione a punteggi (scoring) della propria affidabilità fiscale. Il software assegna precisi benefici a chi ottiene i punteggi più elevati (sopra l'8 su 10), e prescrive invece controlli più severi per chi non raggiunge la sufficienza (6 su 10). Il giudizio si basa su mere presunzioni automatiche di anomalie rispetto a consumi, possedimenti, investimenti e altri comportamenti finanziari. La logica è fornire ai soggetti fiscali un modo per auto-valutare la propria posizione col fisco, così che sia loro possibile colmare qualunque potenziale mancanza — pri-

ma che finisca per attenzionare le autorità, indirizzandole verso ulteriori controlli.

Ma, anche in questo caso, non sono mancati detrattori che hanno sostenuto che il sistema fosse fallato alle fondamenta. Il team 'Dataroom' del Corriere della Sera ha intervistato "molti commercialisti e in diverse regioni italiane", e anche se il metodo utilizzato non consente di dare valenza statistica e/o peso scientifico ai risultati da lei ottenuti, la giornalista investigativa Milena Gabanelli ha concluso che gli errori potrebbero finire per riguardare il 50% dei contribuenti. A seguito dell'introduzione degli ISA "è emerso", ha scritto Gabanelli insieme ad Andrea Marinelli, "che fra il 40 e il 50 per cento dei contribuenti è passato dall'essere "congruo e coerente" nella dichiarazione dei redditi del 2018 a "insufficiente" in quella del 2019, o viceversa" (Corriere della Sera 2019).

Gli errori sarebbero derivati dal fatto che ai contribuenti non è consentito condividere quali siano i precisi fattori che contribuiscono a determinare il loro reddito, agevolando così rappresentazioni errate e fraintendimenti. Basta una maternità, un immobile sfitto o un improvviso — e contingente — aumento delle spese legali per allertare il sistema e trovarsi potenzialmente segnalati per possibili "anomalie", scrive Gabanelli, dato che l'algoritmo è "programmato male", e sarebbe meno flessibile e affidabile perfino degli ampiamente criticati "studi di settore", basati su un metodo di scrutinio fiscale "induttivo".

Mancano tuttavia ancora dati ufficiali al riguardo.

Inoltre, l'idea che si stia cercando di implementare algoritmicamente una sorta di "Grande Fratello fiscale", evidenziata nel rapporto precedente, rappresenta ancora la cornice interpretativa più comunemente adottata nel dibattito pubblico sull'uso di sistemi di ADM in questioni riguardanti il fisco. Ciò è stato particolarmente vero nel caso del nuovo strumento "anonimizzato" per l'analisi predittiva di Big Data dei comportamenti dei contribuenti, che è a sua volta stato introdotto con la legge di bilancio del 2019 (Sole 24 Ore 2019c).

Chiamato "Evasometro Anonimizzato", questo strumento di analisi fiscale include "algoritmi specifici", ha scritto il quotidiano finanziario Il Sole 24 Ore, ed è in grado di condurre verifiche incrociate sui diversi database mantenuti dal fisco italiano, andando a caccia di incongruenze nei pattern di consumo o in qualunque altra operazione finanziaria. Le anomalie riconosciute dagli "occhi digitali" del sistema saranno poi usate per allertare le autorità tributarie, che potranno così convocare gli individui segnalati per ulteriori e più approfonditi controlli.

Questa forma di verifica a guida algoritmica va inoltre ad aggiungersi all'analisi delle "fonti aperte", tra cui "articoli di giornale, siti e social media", da parte dell'Agenzia delle Entrate. Il sistema di controllo è attivo dal 2016, ma è risultato sconosciuto al grande pubblico fino a gennaio 2020. Il sistema è stato predisposto per "acquisire tutti gli elementi che possono essere utili nella conoscenza del contribuente", aiutando così l'Agenzia a valutare la coerenza della posizione fiscale di ogni individuo. (Corriere della Sera 2019b)

Diverse sentenze — per esempio, quelle emesse dalle Corti di Appello di Brescia e Ancona — hanno anche dato dimostrazione di come tale sistema funzioni realmente. Al loro interno, i giudici hanno anche esplicitamente menzionato "documenti estrapolati da Facebook". Un esempio viene dal caso dell'ex marito costretto a pagare le spese di sostentamento all'ex moglie a causa dei post su Facebook in cui mostrava di avere uno stile di vita chiaramente incompatibile con gli 11.000 euro di reddito dichiarati nel corso dell'anno fiscale precedente.

/ Polizia predittiva: accettiamo che diventi la norma?

Gli esempi di polizia predittiva a base di automazione stanno in Italia ampliando rapidamente il loro raggio d'azione. E se nella prima edizione di 'Automating Society' abbiamo mostrato come il software KeyCrime stia venendo utilizzato a Milano, sono le sperimentazioni di un altro sistema algoritmico, chiamato XLAW, a essersi di recente diffuse in molte altre importanti città italiane.

Proprio come KeyCrime, anche il software XLAW è stato sviluppato da un rappresentante delle forze dell'ordine. Il sistema è stato già adottato nelle città di Napoli, Modena, Prato, Salerno, Livorno, Trieste, Trento e Venezia (Business Insider 2019), dove ha portato all'arresto di un uomo di 55 anni accusato di furto (Polizia di Stato 2018). Il sito di informazione giuridica Altalex riporta che il sistema sarebbe risultato accurato nell'87-93% dei casi a Napoli, nel 92-93% a Venezia, e nel 94% a Prato. Al momento, tuttavia, non sono pubblicamente disponibili auditing o fact-checking indipendenti per alcuna di queste sperimentazioni. (Altalex 2018)

XLAW contiene al suo interno una forte componente predittiva. Nel caso di Venezia, per esempio, la Polizia di Stato ha

affermato che "l'84 per cento dei fatti-reato che sono stati tentati o commessi erano stati previsti dal sistema "XLaw". Di conseguenza, anche XLAW, come già KeyCrime, viene di norma raccontato nei media generalisti come un sistema di successo.

E tuttavia, l'unica risposta ufficiale concernente i reali risultati prodotti da Keycrime a Milano nel corso degli anni racconta una storia diversa. Rispondendo a una richiesta di accesso alle informazioni (FOIA) del giornalista di Vice Italy, Riccardo Coluccini, il Dipartimento per le politiche del personale dell'amministrazione civile e per le risorse strumentali e finanziarie del Ministero degli Interni ha affermato di non possedere "eventuali dati e studi sull'impatto del software nella lotta al crimine", nemmeno circa le "variabili tenute in considerazione dal software per effettuare le analisi". (Vice Italy 2018)

Inoltre, l'amministrazione si è detta contraria a fornire piena trasparenza, dato che il software è brevettato e proprietario, e come tale gode di "speciale protezione" in termini di proprietà intellettuale. La Direzione Centrale Anticrimine del Ministero degli Interni lo usa solamente in comodato d'uso gratuito, ha scritto Vice Italy.

Il creatore di XLAW, l'ispettore della Questura di Napoli Elia Lombardo, descrive il software come una soluzione, probabilistica e dotata di algoritmi di machine learning, per scovare tendenze e pattern negli episodi criminali. La logica che ne informa il funzionamento è quella delle "riserve di caccia", un modello di individuazione dei crimini che assume che i reati predatori (per esempio, furti e rapine) siano al contempo "ciclici" e "stanziali". Ciò implica che si possa dedurre — o, più accuratamente, indurre — come si manifesteranno i comportamenti criminali prima ancora che vengano messi in atto, grazie a un'analisi intelligente della storia criminale di un luogo nel tempo. Lombardo afferma di averlo fatto analizzando attentamente vent'anni di dati, utilizzando tutta l'esperienza acquisita sul campo. Il risultato è un sistema, XLAW appunto, che può essere utilizzato per allertare le forze di polizia e inviare pattuglie sulla scena del crimine, ancor prima che venga compiuto. Al contempo, XLAW fornisce agli agenti di polizia precisi dettagli sui potenziali sospetti: "genere, altezza, cittadinanza, caratteristiche distinguibili, e dati biometrici".

Lombardo afferma che il suo approccio sia stato validato dall'analisi "indipendente" di due università a Napoli — la Federico II e la Parthenope — prima di metterlo all'opera.

/ Documenti strategici sull'AI, e perché sono importanti per l'ADM in Italia

Perfino di fronte al proliferare di esempi di sistemi di ADM in ogni aspetto della vita pubblica dei cittadini italiani, le risposte in termini di scelte politiche sono state perlopiù assenti.

L'ipotizzata Commissione parlamentare sull'Etica dell'Intelligenza Artificiale, di cui si era detto nella prima edizione del rapporto, è stata accantonata, mentre le tre principali forze del governo attualmente in carica (Conte II) ve ne hanno preferita una dedicata al tema della disinformazione — e in particolare a un'indagine di quelle che vengono ancora chiamate "fake news", nonostante il termine sia stato ampiamente criticato nella letteratura accademica e dalla società civile. (Valigia Blu 2019)

Nemmeno il centro trans-disciplinare sull'AI proposto, nel suo white paper del 2018, dal Team per la Trasformazione Digitale è stato realizzato.

Tuttavia, tra gennaio e giugno 2019, trenta esperti scelti dal Ministero dello Sviluppo Economico (MISE) hanno stilato un documento per fornire al governo una prima base su cui costruire la propria "Strategia nazionale per l'intelligenza artificiale".

Il rapporto che ne è risultato è stato reso disponibile in consultazione pubblica dal 19 agosto al 13 settembre 2019 (MISE 2019), e la sua versione finale è stata pubblicata il 2 luglio 2020 (MISE 2020).

La strategia italiana afferma di essere saldamente ancorata ai principi dell'antropocentrismo, dell'affidabilità, e della sostenibilità dell'AI, e di porre "un forte accento" sulla necessità di un uso etico dell'AI e sul bisogno di garantire affidabilità tecnica fin dalla fase di concezione di tali sistemi. L'ambizione è nientemeno che stimolare un nuovo "Rinascimento" basato sull'intelligenza artificiale, chiamato "RenAIssance" nel documento finale.

Per indirizzare il paese verso questa direzione, il rapporto elenca 82 proposte, alcune delle quali rilevanti per comprendere come sia concepito l'ADM in Italia. Anche se non esplicitamente menzionati, i sistemi di ADM potrebbero trarre beneficio dall'opera di "mappatura nazionale" degli esperimenti in corso che verrebbe condotta, in accordo con la Strategia, dalla nuova "cabina di regia" interministeriale e

multi-stakeholder dedicata ai molteplici aspetti regolatori, educativi, infrastrutturali e industriali sollevati dall'implementazione di sistemi di AI, così da regolarli, testarli e promuovere ricerca in materia.

Sulla scia della normativa europea, il documento afferma inoltre che la spiegabilità delle decisioni algoritmiche dovrebbe essere garantita tramite forme di trasparenza procedurale, "ogniqualvolta un sistema di AI prenda una decisione che comporta un forte impatto sulla vita personale". Come per il Libro bianco sull'AI della Commissione UE, il grado di trasparenza dovrebbe essere proporzionale ai rischi derivanti dall'uso del sistema di AI in esame.

Secondo la strategia, produttori e distributori di tali sistemi dovrebbero essere inoltre obbligati a informare gli utenti ogniqualvolta stiano interagendo con una macchina, invece che con un essere umano, mentre appropriate sanzioni e rimedi dovrebbero venire predisposte per tutte le decisioni automatizzate che comportino effetti discriminatori o illeciti. Coerentemente con il GDPR, ciò significa che gli utenti devono poter esercitare il loro diritto a sottrarsi a decisioni automatiche, in modo "semplice e immediato".

Pur se ancora in via di definizione, anche l'idea di una "valutazione d'impatto" per una AI "degna di fiducia" ("Trustworthy AI") viene esplicitamente menzionata tra gli strumenti necessari a meglio comprendere i reali rischi derivanti dall'uso di sistemi automatizzati. Sono i loro benefici, tuttavia, a venire insieme fortemente enfatizzati, se si pensa che, secondo la Strategia, soluzioni a base di AI dovrebbero essere adottate per ogni "sfida sociale", e in particolar modo per ottenere una "migliore qualità della vita" — per esempio attraverso l'adozione di AI nella sanità, nell'educazione, e nelle "digital humanities", il campo di studi, in forte crescita, all'intersezione tra analisi del digitale e critica culturale.

L'intelligenza artificiale ricopre un ruolo di primo piano nel documento strategico intitolato '2025. Strategia per l'innovazione tecnologica e la digitalizzazione del paese', steso dal ministero per l'Innovazione tecnologica e la digitalizzazione creato dal secondo (e attuale) governo Conte. Al suo interno, l'AI viene concepita come "al servizio della nazione" e, con un riferimento implicito ai processi di ADM, particolarmente utile nel consentire di ottenere maggiore "efficienza" per procedure amministrative ripetitive, che richiedono un basso grado di discrezionalità.

Al contempo, il documento sostiene che solo un'AI affidabile ed etica contribuirebbe a rispettare il principio costituzionale secondo cui i cittadini hanno diritto a un giusto e tempestivo processo. Non a caso è proprio il sistema giudiziario il primo settore della vita pubblica esplicitamente menzionato nella Strategia per il 2025 come passibile di sperimentazioni. Verrà anche creata una nuova "Alleanza per l'AI sostenibile", un Comitato composto insieme da entità private e istituzioni pubbliche, che avrà il compito di pervenire a uno "Statuto etico giuridico dell'intelligenza artificiale".

"Oltre a fissare un set minimo di principi-guida", l'Alleanza — attraverso il cosiddetto "AI Ethical LAB-EL" — stabilirà "un insieme di regole minime per la qualificazione di soluzioni di intelligenza artificiale destinate al settore pubblico come a quello privato". Secondo la Strategia 2025, il rispetto di queste linee guida potrebbe diventare, nel tempo, l'equivalente di una certificazione "ufficiale" che stabilisce che un certo sistema di AI è sia etico che legale.

Il documento, tuttavia, non contiene menzione della Strategia nazionale per l'AI, e non fornisce alcuna coordinazione tra i due piani ufficiali per una futura Italia fondata sull'intelligenza artificiale.

Da ultimo, ad aprile 2019, il deputato di +Europa, Alessandro Fusacchia, ha lanciato un "Intergruppo parlamentare" sull'AI. (CorCom 2019)

Conclusioni

L'Italia deve essere maggiormente conscia di quanto rapidamente i sistemi che automatizzano processi decisionali si stiano diffondendo in diversi settori dell'amministrazione e della vita pubblica. Tecnologie di riconoscimento facciale e altri sistemi di ADM stanno venendo utilizzati in modo crescente, in assenza di un qualunque dibattito realmente democratico e/o di un significativo dialogo con la società civile e il mondo accademico, preparando così potenzialmente il terreno a futuri utilizzi a loro volta problematici.

Il fallimento dell'algoritmo de La Buona Scuola, ora incastonato in più sentenze, dovrebbe fungere da monito: l'automazione può portare a conseguenze realissime e dolorose quando adottata con troppa impazienza e mal concepita, sia per specifiche categorie sociali (in questo caso, le migliaia di insegnanti assegnati a errata destinazione dalle falle

dell'algoritmo) che per la popolazione in genere (i contribuenti che, in ultima analisi, finiscono per pagare il prezzo dei legittimi ricorsi).

Un principio più generale sembra stare emergendo da questo stato di cose. Per essere davvero benefica e degna di fiducia, così da realizzare quello che la Ministra per l'Innovazione, Paola Pisano, definisce "un nuovo Umanesimo dell'AI e un Rinascimento digitale", l'intelligenza artificiale deve operare insieme a quella umana, aumentandola piuttosto che rimpiazzandola. Se ciò poi significhi che tra gli usi pubblici dell'ADM non possano rientrare decisioni completamente autonome, o se al contrario ciò comporti unicamente una migliore definizione di quali decisioni possono essere completamente automatizzate, è questione ancora controversa. È tuttavia la questione dalla cui risposta dipenderanno, probabilmente, le scelte politiche dell'Italia in tema di ADM — quando ci saranno.

Riferimenti bibliografici

Aeroporti di Roma, Il controllo biometrico, http://www.adr.it/controllo-biometrico#

AgendaDigitale (2019), Giustizia predittiva: così a Brescia l'intelligenza artificiale aiuta a prevedere i tempi delle cause, https://www.agendadigitale.eu/cultura-digitale/giustizia-predittiva-cosi-a-brescia-lintelligenza-artificiale-aiuta-a-prevedere-tempi-e-esiti-delle-cause/

Altalex (2018), X-LAW: la polizia predittiva è realtà, https://www.altalex.com/documents/news/2018/11/28/x-law-la-polizia-predittiva

BVA DOXA (2019), Lavoro, la sfida dei robot e dell'IA, https://www.bva-doxa.com/lavoro-la-sfida-dei-robot-e-dellia/

BresciaOggi (2018), Tempi dei processi, un aiuto dalla "giustizia predittiva", https://www.bresciaoggi.it/territori/citt%C3%A0/tempi-dei-processi-un-aiuto-dalla-giustizia-predittiva-1.6415532

Business Insider (2019), Parla l'ispettore di polizia che ha creato un algoritmo che prevede scippi, rapine, furti. E funziona, da anni, https://it.businessinsider.com/perla-il-dirigente-di-polizia-ha-creato-un-algoritmo-che-prevede-scippi-rapine-furti-e-funziona-da-anni/

Corriere Milano (2019), Linate, riapre l'aeroporto ma il restyling durerà un anno e mezzo, https://milano.corriere.it/notizie/cronaca/19_ottobre_23/linate-riapre-l-aeroporto-ma-restyling-durera-anno-mezzo-5c30c9ee-f55c-11e9-ad25-1b4f9a2d3eb8.shtml

Corriere della Sera (2019), Tasse, ecco come un algoritmo difettoso ti fa pagare più di quanto devi, https://www.corriere.it/dataroom-milena-gabanelli/tasse-isa-ecco-come-algoritmo-difettoso-ti-fa-pagare-piu-quanto-devi/77bbc760-03dd-11ea-a09d-144c1806035c-va.shtml

Corriere della Sera (2019b), Evasione, da social network e big data la nuova caccia ai "furbetti", https://www.corriere.it/economia/tasse/20_gennaio_12/evasione-social-network-big-data-nuova-caccia-furbetti-016742ae-3521-11ea-b1ac-094115aa5a02.shtml

CorCom (2019), Nasce l'interruppe parlamentare sull'intelligenza artificiale, https://www.corrierecomunicazioni.it/digital-economy/nasce-lintergruppo-parlamentare-sullintelligenza-artificiale/

Forbes Italia (2019), Riconoscimento facciale: l'Italia è ultima in Ue per la protezione della privacy, https://forbes.it/2019/10/21/riconoscimento-facciale-litalia-ultima-in-europa-per-protezione-della-privacy/

Giornale Monza (2019), Intelligenza artificiale in corsia: l'esempio dell'ASST di Vimercate, https://giornaledimonza.it/attualita/intelligenza-artificiale-in-corsia-lesempio-dellasst-di-vimercate/

Il Giorno (2019), Vimercate, svolta hi-tech in ospedale: se ti ammali lo prevede l'algoritmo, https://www.ilgiorno.it/monza-brianza/cronaca/ospedale-vimercate-algoritmo-1.4593259

Gonews (2019), Giustizia e tecnologia: l'intelligenza artificiale potrà prevedere le sentenze?, https://www.gonews.it/2019/09/17/giustizia-tecnologia-lintelligenza-artificiale-potra-prevedere-le-sentenze/

Il Messaggero (2020), Coronavirus, il futuro degli aeroporti: intelligenza artificiale e plexiglass, https://www.ilmessaggero.it/roma/senzarete/aeroporto_coronavirus_fiumicino-5176984.html

Ministero per l'Innovazione e la Digitalizzazione (2019), 2025 Strategia per l'innovazione tecnologica e la digitalizzazione del Paese, https://innovazione.gov.it/assets/docs/MID_Book_2025.pdf

MISE (2019), Strategia nazionale per l'intelligenza artificiale, https://www.mise.gov.it/images/stories/documenti/Strategia-Nazionale-Intelligenza-Artificiale-Bozza-Consultazione.pdf

MISE (2020), Proposte per una strategia nazionale per l'intelligenza artificiale, https://www.mise.gov.it/images/stories/documenti/Proposte_per_una_Strategia_italiana_AI.pdf

Mobilita (2019), Aeroporto Fiumicino. Parte il riconoscimento facciale che permetterà di viaggiare senza documenti, https://mobilita.org/2019/11/17/aeroporto-fiumicino-parte-il-riconoscimento-facciale-che-permettera-di-viaggiare-senza-documenti/

Polizia di Stato (2018), Venezia: i reati vengono previsti da un algoritmo, https://www.poliziadistato.it/articolo/225bdd72498c36f619718140

Repubblica (2019), Scuola, trasferimenti di 10 mila docenti lontano da casa. Il Tar: "L'algoritmo impazzito fu contro la Costituzione", https://www.repubblica.it/cronaca/2019/09/17/news/scuola_trasferimenti_di_10mila_docenti_lontano_da_casa_il_tar_l_algoritmo_impazzito_fu_contro_la_costituzione_-236215790/

Repubblica Napoli (2016), All'aeroporto di Napoli arrivano i varchi automatici: controlli in 20 secondi https://napoli.repubblica.it/cronaca/2016/01/19/news/all_aeroporto_di_napoli_arrivano_i_gates_automatici_controlli_in_20_secondi-131604084/

Rete Disarmo (2019), Il 73% dei cittadini di 10 paesi europei vuole la messa al bando dei "killer robots", https://www.disarmo.org/rete/a/47017.html

State Council (2019), ruling n. 2270, April 8, 2019, https://www.avvocatirandogurrieri.it/images/easyblog_articles/9469/CdS-sentenza-n.-2270-dell8-aprile-2019.pdf

Sole 24 Ore (2018), Algoritmi, no grazie! I candidati vogliono un recruiter in carne e ossa, https://www.ilsole24ore.com/art/algoritmi-no-grazie-candidati-vogliono-recruiter-carne-e-ossa-AE5zR8jG

Sole 24 Ore (2019b), Se l'algoritmo prevede l'evoluzione delle malattie e la disponibilità dei letti, https://www.ilsole24ore.com/art/se-l-algoritmo-prevede-l-evoluzione-malattie-e-disponibilita-letti-ACeQkjq

Sole 24 Ore (2019c), Dagli scontrini ai conti bancari, il Fisco lancia la lotta all'evasione con i dati anonimi, https://www.ilsole24ore.com/art/dagli-scontrini-conti-bancari-fisco-lancia-lotta-all-evasione-i-dati-anonimi-AC5wGau

TAR Lazio (2019), ruling n. 10964, section 3-bis, September 13, 2019, http://images.go.wolterskluwer.com/Web/WoltersKluwer/%7B15d38a3f-3bd0-4c31-a437-b325dfa65f35%7D_tar-lazio-roma-sentenza-13-settembre-2019-numero-10964.pdf

TgCom24 (2019), Volare si cambia: con riconoscimento facciale e intelligenza artificiale tornano i liquidi in valigia, https://www.tgcom24.mediaset.it/tgtech/volare-si-cambia-con-riconoscimento-facciale-e-intelligenza-artificiale-tornano-i-liquidi-in-valigia_11228327-201902a.shtml

Valigia Blu (2019), Perché diciamo no a una Commissione parlamentare d'inchiesta sulle "fake news", https://www.valigiablu.it/commissione-inchiesta-fake-news/

Vice Italy (2018), La polizia predittiva è diventata realtà in Italia e non ce ne siamo accorti, https://www.vice.com/it/article/pa5apm/polizia-predittiva-italia-lombardi-xlaw-prevedere-crimini-algoritmi

Team

/ Beate Autering

Graphic designer e layout artist

Beate Autering è una graphic designer freelance. Si è laureata in design, e gestisce lo studio beworx. Crea progetti, grafiche e illustrazioni, e fornisce inoltre servizi di editing di immagini e post-produzione.

/ Fabio Chiusi

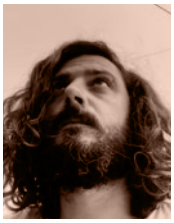
Editor del rapporto, autore della storia giornalistica e del capitolo di ricerca sull'Italia, così come dell'introduzione e del capitolo sull'Unione Europea

Photo: Julia Bornkessel

Fabio Chiusi lavora ad AlgorithmWatch come co-editor e project manager dell'edizione 2020 del rapporto Automating Society. Dopo un decennio nel giornalismo tecnologico, è stato consulente e ricercatore in tema di dati e politica (Tactical Tech) e AI nel giornalismo (Polis LSE). Ha coordinato il rapporto "Persuasori Social" sulla regolamentazione delle campagne politiche sui social media per il Progetto PuntoZero, e ha lavorato come responsabile politiche tecnologiche del Questore della Camera dei Deputati del Parlamento italiano, durante l'attuale legislatura. Fabio è fellow del Nexa Center for Internet & Society di Torino, e docente a contratto all'Università di San Marino, dove insegna Giornalismo e nuovi media ed Editoria e media digitali. È autore di diversi saggi su tecnologia e società, di cui il più recente è 'Io non sono qui. Visioni e inquietudini da un futuro presente' (DeA Planeta, 2018), tradotto in polacco e cinese. Scrive di politiche tecnologiche sul blog collettivo Valigia Blu.

/ Samuel Daveti

Fumettista



Samuel Daveti nasce nel 1983 a Grosseto. Socio fondatore dell'Associazione Culturale Double Shot, per il mercato francese sceneggia l'albo Akron Le guerrier (Soleil, 2009) ed è curatore del volume antologico Fascia Protetta (Double Shot, 2009). Nel 2011 è uno dei fondatori di Mam-maiuto, collettivo di autoproduzione. Sceneggiatore di Un Lungo Cammino (Mammaiuto, 2014; Shockdom, 2017), che diventerà un film per la casa di produzione cinematografica Brandon Box, nel 2018 scrive I Tre Cani per i disegni di Laura Camelli, storia vincitrice del Premio Micheluzzi al Napoli Comicon 2018 e del premio Boscarato al Treviso Comic Book Festival come Miglior Web-comic. Nel 2020 sceneggia Inerti per i disegni di Francesco Rossi, storia vincitrice del premio Bo-scarato al Treviso Comic Book Festival come miglior Webcomic.

/ Sarah Fischer

Editor del rapporto

Sarah Fischer è project manager per il progetto "Ethics of Algorithms" di Bertelsmann Stiftung, per il quale è principalmente responsabile degli studi scientifici. In precedenza, ha lavorato come postdoc fellow nel programma specialistico "Trust and Communication in a Digitalized World" all'Università di Münster, dove si è concentrata sul tema della fiducia nei motori di ricerca. Nello stesso gruppo di addestramento alla ricerca, Sarah ha ottenuto il dottorato con una tesi sulla fi-ducia nei servizi sanitari online. Ha studiato scienze della comunicazione alla Friedrich Schiller University di Jena, ed è co-autrice dei paper "Where Machines can err. Sources of error and re-sponsibilities in processes of algorithmic decision making" e "What Germany know and believes about algorithms".

/ Leonard Haas

Ulteriore revisione dei testi

Leonard Haas lavora come assistente ricercatore ad AlgorithmWatch. Tra le altre cose, è stato responsabile della concezione, dell'implementazione e del mantenimento dell'AI Ethics Guidelines Global Inventory. È studente in un master nel campo delle scienze sociale alla Humboldt University di Berlino e ha ottenuto dall'Università di Lipsia due lauree di primo livello, in Digital Humanities e Scienze Politiche. La sua attività di ricerca si concentra sull'automazione del lavoro e della gover-nance. In aggiunta, Leonard si occupa di politiche dei dati di interesse pubblico e di lotte dei lavoratori del settore tecnologico.

/ Lorenzo Palloni

Fumettista

Lorenzo Palloni, fumettista, classe 1987. Fra i fondatori di Mammaiuto, è autore de La lupa (Sal-dapress, 2019); Mooned (Shockdom, 2017); Scary Allan Crow (Edizioni Inkiostro, 2017) è sceneg-giatore di Desolation Club con Vittoria Macioci (Saldapress, 2019); Emma Wrong con Laura Gu-glielmo (Editions Akileos, 2019); Instantly Elsewhere (Shockdom, 2018) e Terranera (Feltrinelli, 2020), entrambi con Martoz. È stato ospite della "Maison Des Auteurs" di Angoulême e ha vinto due premi Boscarato come "Miglior Sceneggiatore Italiano" e, nel 2019, il premio Gran Guinigi come "Miglior Sceneggiatore". Al momento sta lavorando a libri di prossima uscita per il mercato francese e italiano, ed è docente di Sceneggiatura e Storytelling alla "Scuola Internazionale di Comics" nelle sedi di Firenze e Reggio Emilia.

/ Kristina Penner

Autrice del capitolo sull'Unione Europea

Photo: Julia Bornkessel

Kristina Penner è executive advisor ad AlgoritmWatch. I suoi interessi di ricerca includono l'ADM nei sistemi di welfare, le forme di scoring sociale, e gli impatti dell'ADM sulla società, così come la sostenibilità delle nuove tecnologie da un punto di vista olistico. La sua analisi dei sistemi di ge-stione dei confini dell'UE è una prosecuzione della sua precedente esperienza di ricerca e consu-lenza sulle leggi in materia di asilo. Altre precedenti esperienze di Kristina includono progetti sull'uso dei media nella società civile e nel giornalismo "conflict-sensitive", così come il coinvol-gimento come stakeholder nel processo di pace nelle Filippine. Kristina ha ottenuto un master in studi internazionali/ricerca su pace e conflitto alla Goethe University di Francoforte.

/ Alessio Ravazzani

Fumettista

Alessio Ravazzani nasce a Cuggiono (MI) l'11 maggio del 1975. Si diploma al Liceo artistico Can-diani di Busto Arsizio nel 1993 e nel 2007 si diploma alla Scuola internazionale di Comics di Firen-ze. Grafico editoriale, fumettista e illustratore, si approccia al lettering come professionista nel 2010 con delle collaborazioni con le Edizioni BD e la sua divisione J-POP. Nel 2011 comincia la collaborazione con lo studio della divisione Lion della casa editrice RW, proprietaria dei diritti della DC Comic per l'Italia; quattro anni più tardi, quella con Saldapress, con la francese Kazé e l'americana Viz media. Nel 2016 si aggiungono la Symmaceo e lo studio RAM: ambedue partner della casa editrice Panini comics. Nel 2018 comincia a collaborare con Coconino-Fandango. Dal 2011 è membro del collettivo di autori Mammaiuto con cui collabora come disegnatore e grafico. Attualmente abita e lavora alla Spezia.

/ Matthias Spielkamp

Editor del rapporto

Photo: Julia Bornkessel

Matthias Spielkamp è co-fondatore e direttore esecutivo di AlgorithmWatch. Ha testimoniato di fronte a diverse commissioni del Parlamento federale tedesco in tema di AI e automazione. Mat-thias siede nel board direttivo della sezione per la Germania di Reporters Without Borders e nei comitati consultivi di Stiftung Warentest e del Whistleblower Network. È stato fellow di ZEIT Stif-tung, Stiftung Mercator e dell'American Council on Germany. Matthias ha anche fondato il maga-zine online mobilsicher.de, che si occupa di sicurezza dei telefonini, e ha un pubblico di oltre 170.000 lettori mensili. Ha scritto e curato libri sul giornalismo digitale e la governance di Internet, ed è stato tra i 15 architetti del nostro futuro "data-driven" scelti da Silicon Republic. Matthias ha ottenuto un master in giornalismo all'Università del Colorado, a Boulder, e un master in filosofia all'Università Libera di Berlino.

/ Tiger Stangl

Graphic designer aggiunta e layout artist

Tiger Stangl è una designer che lavora a Berlino su progetti di design editoriale di beworx, Friedrich Ebert Stiftung, Buske Verlag, UNESCO Welterebe Deutschalnd e.V., Sehstern Agency, Rights Lab, e Landersspracheninstitut.

/ Marc Thümmler

Coordinatore della pubblicazione

Photo: Julia Bornkessel

Marc Thümmler è responsabile delle relazioni pubbliche e della comunicazione ad AlgorithmWatch. Ha un master in media studies, e ha lavorato come produttore ed editor in una compagnia cinematografica, oltre a gestire progetti per la Deutsche Kinemathek e per l'organizzazione della società civile Gesicht Zeigen. In aggiunta ai suoi principali incarichi, Marc è stato anche coinvolto nelle campagne di crowdfunding e crowdsourcing per OpenSCHUFA, e ha inoltre coordinato la prima edizione del rapporto Automating Society, pubblicata nel 2019.

ORGANIZZAZIONI

/ AlgorithmWatch

AlgorithmWatch è una organizzazione no profit di ricerca e advocacy che si dedica a studiare, comprendere e analizzare i sistemi di decision-making algoritmico/automatizzato (ADM) e il loro impatto sulla società. Se un ricorso accorto ai sistemi di ADM può arrecare benefici tanto a singoli individui quanto a intere comunità, il loro uso si accompagna anche a grossi rischi. Per difendere l'autonomia umana e i diritti fondamentali, e massimizzare il bene pubblico, riteniamo cruciale sottoporre i sistemi di ADM a scrutinio democratico. L'uso di sistemi di ADM che abbiano un significativo impatto sui diritti individuali e collettivi non dovrebbe unicamente essere reso pubblico in modi chiari e accessibili, ma dovrebbe anche consentire a ogni singolo individuo di comprendere come si sia pervenuti a una decisione, e di contestarla se necessario. Per questo aiutiamo i cittadini a meglio comprendere i sistemi di ADM, così come a svilupparne nuove modalità di governance — attraverso un misto di tecnologie, regole e adeguati organi di controllo. È questo il nostro modo di dare un contributo verso una società più equa e inclusiva, oltre che per massimizare i benefici dei sistemi di ADM per la società tutta.

https://algorithmwatch.org/en/

ALGORITHM WATCH

/ Bertelsmann Stiftung

L'obiettivo di Bertelsmann Stiftung è promuovere l'inclusione sociale per tutti. L'impegno è a perseguire lo scopo attraverso programmi mirati al miglioramento dell'istruzione, a informare la democrazia, far progredire la società, promuovere la salute, rivitalizzare la cultura e rinforzare le economie. Tramite le sue attività, Bertelsmann Stiftung si propone di incoraggiare i cittadini a contribuire al bene pubblico. Fondata nel 1977 da Reinhard Mohn, la fondazione no profit detiene la maggioranza delle azioni di Bertelsmann SE & Co. KGaA. Bertelsmann Stiftung è una fondazione privata e indipendente. Con il suo progetto "Etica degli Algoritmi", sta fornendo uno sguardo approfondito sulle conseguenze sociali dei processi decisionali algoritmici, allo scopo di assicurare che tali sistemi siano utilizzati per il bene collettivo. La fondazione intende così contribuire a informare e promuovere sistemi algoritmici che facilitino l'inclusione sociale. Ciò comporta impegnarsi in ciò che è meglio per la società, piuttosto che in ciò che è tecnicamente possibile — così che le decisioni prese con le macchine siano al servizio del genere umano nel modo più proficuo.

https://www.bertelsmann-stiftung.de/en

BertelsmannStiftung