LA *PREDICTIVE POLICING* NEL REGOLAMENTO EUROPEO SULL'INTELLIGENZA ARTIFICIALE

di Elisabetta Pietrocarlo
(Assegnista di ricerca in diritto penale, Università Luiss)

Sommario: 1. Polizia predittiva e *AI Act*: una premessa. – 2. La c.d. *predictive policing*: origini e tipologie di sistemi. – 3. I profili critici. – 4. L'inquadramento della polizia predittiva nell'*AI Act*: un percorso tortuoso – 5. La disciplina della polizia predittiva nell'*AI Act*. – 6. Osservazioni conclusive e prospettive future.

1. L'uso di algoritmi predittivi da parte delle forze dell'ordine è da diverso tempo oramai al centro del dibattito pubblico e scientifico a livello globale. Grazie all'esperienza principe degli Stati Uniti, anche in Europa è progressivamente maturata la consapevolezza circa le straordinarie potenzialità analitiche dei sistemi algoritmici, specie se basati sull'intelligenza artificiale (AI o IA), volti a elaborare previsioni utili a orientare in modo più preciso e razionale l'operato degli agenti di polizia: il riferimento è, come noto, alla c.d. *predictive policing*[1]. Tuttavia, proprio la prassi sviluppatasi nei sistemi d'oltreoceano ha messo a nudo le problematicità legate al ricorso a siffatti

[1] Nella (sterminata) letteratura americana, v. *ex multiis* A.G. Ferguson, T*he Rise of Big Data Policing. Surveillance, Race, and The Future of Law Enforcement,* New York 2017; Id., *Policing Predictive Policing*, in *Washington Law Review* 2017, vol. 94, no. 5, 1109 ss.; Id., *Illuminating Black Data Policing*, in *Ohio State Journal of Criminal Law* 2018, 15, 503 ss.; E. Joh, *Artificial Intelligence and Policing: First Questions*, in *Seattle University Law Review* 2018, 41, 1139 ss.; Id., *Feeding the Machine: Policing, Crime Data, & Algorithms*, in *William & Mary Bill of Rights Journal* 2017, 26, 287 ss.; J. Bachner, *Predictive Policing: Preventing Crime with Data and Analytics*, Washington DC 2013; W.L. Perry, B. McInnis, C.C. Price, S.C. Smith, J.S. Hollywood, *Predictive Policing. The Role of Crime Forecasting in Law Enforcement Operations*, Washington D.C. 2013; A. Selbst, *Disparate Impact in Big Data Policing*, in *Georgia Law Review* 2017, vol. 52, 109 ss. Nella dottrina italiana, v., senza pretesa di esaustività, L. Algeri, *Intelligenza artificiale e polizia predittiva*, in *DPP* 2021, 724 ss.; F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in www.dirittopenaleuomo.org, 29.9.2019, 10 ss.; A. Bonfanti, *"Big data" e polizia predittiva: riflessioni in tema di protezione del diritto alla "privacy" e dei dati personali*, in *MediaLaws* 2018, 3, 10 ss.; G. Contissa, G. Lasagni, G. Sartor, *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in *DInternet* 2019, 4, 620 s.; V. Manes, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *Discrimen*, 15.2.2020, 6 s.; C. Parodi, V. Sellaroli, *Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco*, in *DPenCont* 2019, 6, 55 ss.; P. Severino, *Intelligenza artificiale e diritto penale*, in *Intelligenza artificiale: il diritto, i diritti, l'etica*, a cura di U. Ruffolo, Milano 2020, 540 ss.

strumenti, facendo chiaramente emergere, anche nel contesto europeo, l'esigenza di regolamentazione del fenomeno in esame[2].

Una risposta è oggi offerta dal c.d. *AI Act*[3], il primo tentativo di disciplinare le diverse modalità di impiego degli strumenti in questione, incluse quelle riguardanti il delicato settore della giustizia penale, secondo un approccio di carattere orizzontale[4] e basato sul rischio[5]. L'obiettivo è quello di introdurre regole armonizzate dirette ai diversi soggetti pubblici e privati che sviluppano, distribuiscono, utilizzano, etc. "sistemi di AI"[6] nel territorio dell'Unione, indipendentemente dalla loro ubicazione[7].

---

[2] Basti qui ricordare che, nelle more dell'adozione dell'*AI Act*, è stata adottata la Risoluzione del Parlamento europeo del 6.10.2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale (2020/2016(INI)), con la quale – oltre a chiedere una moratoria agli Stati membri rispetto all'utilizzo dei sistemi di identificazione biometrica – l'organo rappresentativo esprimeva profonde preoccupazioni, proprio richiamando l'esperienza negativa degli Stati Uniti riguardo alla polizia predittiva, in quella sede opponendosi «all'utilizzo dell'IA da parte delle autorità di contrasto per fare previsioni sui comportamenti degli individui o di gruppi sulla base di dati storici e condotte precedenti, dell'appartenenza a un gruppo, l'ubicazione o qualunque altra caratteristica al fine di identificare le persone che potrebbero commettere un reato» (v. paragrafo 24). Per un'analisi della Risoluzione, v. A. Valsecchi, *L'intelligenza artificiale nel diritto penale: la Risoluzione del Parlamento europeo del 6 ottobre 2021*, in *MediaLaws*, 1.2.2022.

[3] V. reg. 1689/2024/UE del Parlamento europeo e del Consiglio del 13.6.2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale), pubblicato sulla Gazzetta ufficiale dell'Unione europea il 12.7.2024.

[4] Ciò significa che le regole ivi stabilite dovrebbero applicarsi in tutti i settori e senza pregiudicare la normativa vigente dell'Unione (ad esempio, in materia di protezione dei dati, diritti fondamentali, etc.) a cui il Regolamento è complementare. V. in tal senso il Considerando n. 9.

[5] Per alcune considerazioni sulla scelta del *risk-based approach* nell'*AI Act*, v. C. Novelli, *L'*Artificial Intelligence Act *Europeo: alcune questioni di implementazione*, in *Federalismi.it* 2024, 2, 97 ss.

[6] Ai sensi dell'art. 3, punto 1) reg. 1689/2024/UE, per "sistema di IA" si intende «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali». Una definizione – almeno nelle intenzioni – *future-proof*, come osserva il Considerando n. 12 che ne sottolinea per l'appunto la flessibilità alla celere evoluzione tecnologica. Per alcune considerazioni critiche al riguardo, v. per tutti M. Colacurci, *Quale diritto penale dell'IA?: alcune riflessioni a partire dalla proposta di regolamento dell'Unione Europea*, in *Jus* 2023, 363. Sulla complessa ricostruzione del significato del termine "intelligenza artificiale", v. B. Panattoni, *Intelligenza artificiale: le sfide per il diritto penale nel passaggio dall'automazione tecnologica all'autonomia artificiale*, in *DInf* 2021, 2, 317 ss.; A. Giannini, *Criminal Behavior and Accountability of Artificial Intelligence Systems*, The Hague 2023, 5 ss.

[7] V. art. 2 reg. 1689/2024/UE nonché Considerando nn. 21, 22, 23. Il n. 22 specifica che l'applicabilità dell'*AI Act* anche ai fornitori e ai *deployer* di sistemi di IA stabiliti in un Paese terzo è opportuna allorché il relativo *output* sia destinato a essere utilizzato nell'Unione. L'obiettivo è quello di scongiurare pratiche elusive e, di conseguenza, garantire una efficace protezione delle persone che si trovano nel territorio eurounitario. V. L. Floridi, *The European Legislation on AI: a Brief Analysis of its Philosophical Approach*, in *Philosophy & Technology* 2021, 217, che osserva come ciò estenderà il c.d. *Brussels effect*, ovverosia l'ampliamento *de facto* delle leggi dell'Unione oltre i suoi confini geografici attraverso i meccanismi di mercato, in quanto anche le imprese extra-UE finiranno per conformarsi al Regolamento; v. altresì M. Colacurci, *op. cit.*, 359. In generale sul tema, v. A.

Nella prospettiva di rafforzare il mercato interno e di promuovere un approccio antropocentrico dell'AI nonché la *leadership* mondiale dell'Unione nello sviluppo di un'AI sicura, affidabile ed etica[8], l'*AI Act* stabilisce una serie di prescrizioni la cui tipologia e il cui contenuto sono plasmati, secondo un criterio di proporzionalità, sulla base dell'intensità e della portata dei rischi che i sistemi di AI possono generare[9]. Il Regolamento identifica pertanto tre categorie al cui interno classificare i diversi *AI systems*, vietando le pratiche che presentano un livello di rischio inaccettabile[10] – come tale incompatibile con il quadro di valori e principi condivisi a livello europeo –; sottoponendo i sistemi ad alto rischio – quelli cioè che possono avere «un impatto nocivo significativo sulla salute, la sicurezza e i diritti fondamentali delle persone nell'Unione»[11] – al rispetto di determinati requisiti nonché all'osservanza, da parte dei pertinenti operatori, di specifici obblighi; imponendo, infine, obblighi di trasparenza per i sistemi caratterizzati da un rischio limitato per i diritti e le libertà individuali[12]. Sono poi previste regole specifiche per i sistemi di AI con finalità generali[13], mentre non rientrano nel campo di applicazione del Regolamento quelli che presentano rischi minimi o assenti, i quali sono permessi senza restrizioni, pur incoraggiandosi l'adozione di codici di condotta volontari.

La particolare rilevanza delle scelte compiute a Bruxelles impone di confrontarsi con i loro riflessi sulle diverse applicazioni di polizia predittiva. Che si tratti di un tema "caldo" e di sicura attualità è dimostrato non solo dalla crescente attenzione riservata a tali sistemi, come si ricordava, nella riflessione scientifica e nel confronto pubblico, ma anche dal "moto ondivago" che ha caratterizzato, come vedremo, l'*iter* legislativo

Bradford, *The Brussels Effect: How the European Union Rules the World*, Oxford 2020.

[8] V. art. 1 reg. 1689/2024/UE che esplicita il relativo scopo.

[9] La ragione di una simile scelta è evidentemente quella di contenere la "pressione normativa" al minimo indispensabile per fronteggiare i rischi l'AI pone. Sull'approccio prescelto dal legislatore europeo, v. G. Lo Sapio, *Intelligenza artificiale: rischi, modelli regolatori, metafore*, in *Federalismi.it* 2022, 27, 238 ss.

[10] V. art. 5 reg. 1689/2024/UE.

[11] Così il Considerando n. 46. Si tratta dei sistemi che soddisfano entrambe le condizioni di cui all'art. 6 – «a) il sistema di IA è destinato a essere utilizzato come componente di sicurezza di un prodotto, o il sistema di IA è esso stesso un prodotto, disciplinato dalla normativa di armonizzazione dell'Unione elencata nell'allegato I; b) il prodotto, il cui componente di sicurezza a norma della lettera *a* è il sistema di IA, o il sistema di IA stesso in quanto prodotto, è soggetto a una valutazione della conformità da parte di terzi ai fini dell'immissione sul mercato o della messa in servizio di tale prodotto ai sensi della normativa di armonizzazione dell'Unione elencata nell' allegato I» – e di quelli indicati nell'Allegato III.

[12] V. art. 50 reg. 1689/2024/UE.

[13] V. Capo V reg. 1689/2024/UE (artt. 51 ss.).

europeo rispetto alla posizione da assumere proprio riguardo alla *predictive policing*, di cui, peraltro, si registrano alcuni esemplari anche nell'ordinamento italiano.

Nel presente lavoro muoveremo da un'analisi della prassi, essenzialmente nordamericana, relativa all'utilizzo di algoritmi predittivi ad opera delle forze dell'ordine[14] e delle criticità registratesi[15], per poi comprendere come inquadrare le diverse tipologie di *software* di cui si tratta nell'architettura dell'*AI Act*[16], con i relativi riflessi in punto di disciplina applicabile[17].

Possiamo sin da subito anticipare che il Regolamento sembra ammettere a certe condizioni il ricorso agli strumenti in esame e, dunque, ci soffermeremmo nell'ultima parte dello scritto sulle prospettive di utilizzo nel nostro ordinamento[18].

2. Letteralmente definita come «*the use of data and analytics to predict crime*»[19], la polizia predittiva si riferisce, in linea generale, all'impiego di tecniche di analisi dei dati volte a elaborare previsioni relative ai luoghi in cui è altamente probabile che siano commessi reati (c.d. *crime hot spots*) ovvero riguardanti i soggetti più a rischio di divenire autori (*potential offenders*) o vittime di reati.

L'interesse verso la *predictive policing* è cresciuto notevolmente, anche a livello europeo, negli ultimi anni in coincidenza con lo sviluppo e la diffusione dell'AI; già prima però del ricorso alle moderne tecniche di *machine learning* e della scoperta del "nuovo petrolio" – i c.d. *big data*[20] – i dipartimenti di polizia americani, a partire dagli anni '90, avevano iniziato a impiegare algoritmi per assicurare che il comportamento degli agenti fosse improntato a maggiore oggettività e risultasse più facilmente verificabile[21]. L'evoluzione delle tecnologiche digitali e la loro capacità di prestazione sotto il profilo della migliore allocazione delle risorse e della più precisa identificazione delle aree urbane o dei soggetti maggiormente "a rischio" hanno condotto le forze di polizia a mutare le proprie strategie: da un approccio esclusivamente reattivo si è virato

[14] V. *infra* § 2.
[15] V. *infra* § 3.
[16] V. *infra* § 4.
[17] V. *infra* § 5.
[18] V. *infra* § 6.
[19] La definizione è di A. Selbst, *op. cit.*, 114.
[20] Sul tema, v. *ex multiis*, F. Di Porto, *La rivoluzione* big data*. Un'introduzione*, in *Concorrenza e mercato* 2016, 5 ss.; A. Ottolia, Big Data *e innovazione computazionale*, Torino 2018; V. Zeno-Zencovich, *Dati, grandi dati, dati granulari e la nuova epistemologia del giurista*, in *MediaLaws* 2018, 2, 32 ss.
[21] Per una ricostruzione delle origini della polizia predittiva, v. A.G. Ferguson, *The Rise of Big Data Policing*, cit., 29 ss.

verso uno prevalentemente proattivo, secondo cui l'intervento delle forze dell'ordine precede – e prescinde da(l) – l'attività criminale, al fine di prevenirla[22].

Chiariti l'obiettivo e i tratti di fondo della *predictive policing*, è ora possibile addentrarsi nell'analisi delle diverse tipologie di *software* emerse nella prassi, seguendo la principale classificazione basata sul diverso oggetto della predizione (luogo/persona). In quest'ottica si distinguono i *place-based systems* e i *person-based systems*[23]. All'interno di quest'ultimo insieme è da taluno[24] individuato un sottogruppo di sistemi denominati *suspect-based*, i quali delineano il profilo del possibile autore di un determinato reato o, più spesso, di una serialità criminale. Invero, come si dirà, proprio l'esperienza italiana vanta un'interessante applicazione di questa tipologia di *tools*.

I primi rappresentano i sistemi più risalenti ma tuttora maggiormente diffusi, probabilmente perché, come avremo modo di evidenziare, si sono rivelati decisamente meno problematici dei secondi.

Essi mirano a individuare le aree geografiche nelle quali, in un determinato arco temporale, è altamente probabile siano commessi reati, cosicché gli agenti possano intensificare i controlli con l'obiettivo di prevenire l'attività criminale ovvero di arrestare in flagranza i relativi autori[25]. L'idea di fondo è quella secondo cui l'aumento delle pattuglie negli *hot spot* possa avere un effetto deterrente sui potenziali criminali[26].

Vediamo di analizzare più nel dettaglio il meccanismo di funzionamento dei sistemi in esame. Partiamo da uno dei primi e più diffusi esemplari di polizia predittiva negli USA, ossia *PredPol*[27], un algoritmo predittivo sviluppato sullo schema di quello

[22] I. Mugari, E.E. Obioha, *Predictive Policing and Crime Control in The United States of America and Europe: Trends in a Decade of Research and the Future of Predictive Policing*, in *Social Sciences* 2021, 1; A.G. Ferguson, *Policing Predictive Policing*, cit., 1137.

[23] Tale ripartizione è seguita, ad esempio, da A.G. Ferguson, *The Rise of Big Data Policing*, cit., *passim*, nonché in sostanza da L. Sommerer, *Self-imposed Algorithmic Thoughtlessness and the Automation of Crime Control*, Baden-Baden 2022, 30, che distingue tra sistemi *person-based* e *location-based*. Per una compiuta ricognizione delle diverse classificazioni, si rinvia a v. I. Mugari, E.E. Obioha, *op. cit.*, 5 s. V. in particolare, oltre a quella da noi prescelta, quelle proposte rispettivamente da W.L. Perry, B. McInnis, C.C. Price, S.C. Smith, J.S. Hollywood, *op. cit.*, 36 e da J. Bachner, *op. cit.*, 86 ss.

[24] A. Selbst, *op. cit.*, 139 ss.

[25] A.G. Ferguson, *The Rise of Big Data Policing*, cit., 63.

[26] A.G. Ferguson, *Predictive Policing Theory*, in *Washington College of Law Research Paper* 2020, 10, 494; L. Sommerer, *op. cit.*, 30.

[27] *PredPol* (oggi *Geolitica*) è stato sviluppato grazie a una collaborazione tra alcuni ricercatori dell'UCLA e il Dipartimento di Polizia di Los Angeles. Per un inquadramento completo del sistema in discorso, v. A.G. Ferguson, *Predictive Policing Theory*, cit., 494 s. nonché E. Silverman, *AI and the Administration of Justice in the*

impiegato per misurare le scosse sismiche. Esso si basa sul *c.d. near repeat effect*[28] – ossia sull'idea secondo cui alcuni *property crime*, al pari dei terremoti, sono seguiti dalla commissione di reati simili nelle aree limitrofe. Pertanto, esaminando i dati storici relativi alle forme di criminalità considerate[29], il sistema rileva anzitutto, in relazione a tre variabili (tipologia, data e orario, luogo di commissione), i *patterns* (ovverosia le corrispondenze) tra i reati precedentemente commessi; successivamente, mediante una raffigurazione su una mappa digitale a disposizione dei *police officers*, sono selezionate le aree urbane (ciascuna di estensione pari a 500 metri quadrati per 500) in cui è probabile si verifichino altrettanti episodi criminosi.

Volgendo adesso lo sguardo al contesto europeo, in Germania ci si imbatte in un sistema analogo, *PreCops*, anch'esso basato sul *near repeat effect* e incentrato, del pari, su alcuni *property crime*[30].

Come si accennava in precedenza, le applicazioni di *predictive policing* si sono in seguito evolute [31], avuto riguardo sia al tipo di reati oggetto di predizione – estendendosi, ad esempio, ai *violent crime* – sia alla teoria di fondo su cui si basa l'analisi algoritmica – il riferimento è al *Risk Terrain Modeling*[32] – sia, infine, alla natura dei dati forniti al sistema per elaborare la previsione – ai dati storici sui reati si sono aggiunti, addirittura, quelli relativi a condizioni metereologiche, eventi sportivi etc.[33].

United States of America: Predictive Policing and Predictive Justice, in *eRevue Internationale de Droit Pénal* 2023, 2, 6 s.; nella letteratura italiana, v. F. Basile, *Intelligenza artificiale e diritto penale*, cit., 11 s.

[28] Per un'illustrazione del fondamento teorico del *near repeat effect* – individuato nelle *theories* della *rational choice* e della *routine activity* – e per ulteriori riferimenti bibliografi stranieri sul punto, sia consentito rinviare a E. Pietrocarlo, Predictive Policing*: criticità e prospettive dei sistemi di identificazione dei potenziali criminali*, in *DPenCont* 2023, 2, 117 s.

[29] *PredPol* riguarda in particolare analisi predittive relative ai seguenti reati: *burglary*, *automobile theft* e *theft from automobiles* (rispettivamente, furto con scasso, furto d'auto e furto di oggetti dall'automobile).

[30] In particolare, sui *residential bulglary*. In generale, su *PreCobs*, v. J. Sprenger, D. Brodowski, *Predictive Policing in Germany*, in *Revue Internationale de Droit Pénal* 2023, 2, 118 s.

[31] Ripercorre tale processo evolutivo A.G. Ferguson, *The Rise of Biga Data Policing*, cit., 66 ss.

[32] Si tratta di un modello che identifica le aree urbane maggiormente esposte al pericolo di commissione di reati sulla base della rilevazione di fattori di rischio (*environmental crime drivers*) che incentivano la criminalità. In altre parole, considerando non solo i dati relativi a precedenti reati bensì tutti quelli collegabili al loro verificarsi (es. le caratteristiche infrastrutturali di una specifica zona, stazioni della metro, luoghi di ritrovo della movida notturna, etc.), RTM seleziona gli *spot* ove, alla luce della presenza di più fattori di rischio, è più probabile che siano commesse attività criminose. V. J.M. Caplan, L.W. Kennedy, *Risk Terrain Modeling: Crime Prediction and Risk Reduction*, Oakland 2016; tra gli autori italiani che hanno analizzato questo strumento, v. F. Basile, *Intelligenza artificiale e diritto penale*, cit., 11.

[33] Nel panorama americano, un plastico esempio di un simile avanzamento è senz'altro costituito da *HunchLab*, un algoritmo basato sul *machine learning* sviluppato dalla società Azavea che poggia su un processo analitico articolato in due fasi: la prima, in cui si considerano i dati relativi ai reati già commessi; la seconda, in cui si includono altresì nell'analisi i c.d. *non-crime datasets*, e cioè dati relativi a fattori sociali, economici o addirittura

Si colloca in questo ambito il *software* italiano *X-Law*[34] che utilizza tecniche di *machine learning* e analizza non solo i dati relativi allo storico dei reati predatori urbani (cioè furti, rapine, etc.) ma prende anche in esame fattori socio-economici, demografici, etc., raffigurando digitalmente, al pari di *PredPol*, le aree maggiormente a rischio-reato in un determinato lasso temporale.

Su un diverso piano si situano i sistemi *person-based*, che sono diretti all'identificazione dei potenziali autori o vittime di reati[35]. Il retroterra teorico di questi strumenti è rappresentato dalla c.d. *focused deterrence*[36], secondo cui gli episodi di violenza sono riferibili solo a una limitata quota della popolazione, con il riflesso che, per ridurre il tasso di criminalità, è fondamentale intervenire in modo mirato sui soggetti in questione.

Anche in questo caso, l'analisi di alcune applicazioni concrete ci consentirà di comprendere il funzionamento di questi sistemi, per poi esaminare le loro principali criticità[37].

Un primo sistema è quello implementato a Kansas City (Missouri) che si caratterizza per un procedimento articolato in quattro fasi, il cui obiettivo consiste nell'identificare i (soli) soggetti più a rischio di divenire autori (e non anche vittime) di reati[38].

Nel segmento iniziale, si elabora una lista di *target offenders* contenente i sospettati di omicidi, conflitti a fuoco o altre gravi aggressioni; nella seconda fase sono passati in rassegna i contatti formali della polizia con i soggetti in questione al fine di risalire ai loro *associates* (si tratta, ad esempio, di coloro che sono stati arrestati o fermati con l'*initial offender*); nel terzo passaggio si individuano i soggetti a loro volta collegati a questi ultimi, ovverosia agli *associates* dell'*offender* principale. Ciò consente di creare

---

metereologici. Successivamente, il sistema non solo identifica gli *hot spot*, ma suggerisce anche agli agenti la specifica strategia da adottare sul campo. Ulteriori dettagli in A.G. Ferguson, *Predictive Policing Theory*, cit., 496 s.; E. Silverman, *op. cit.*, 7 s., nonché nel documento predisposto dalla stessa Azavea, *HunchLab: Under the Hood*, 2015, consultabile su http://blog.pilpul.me/files/2015/09/HunchLab-Under-the-Hood.pdf.

[34] Si tratta di uno strumento ideato da un esponente della Questura di Napoli poi acquisito e ulteriormente sviluppato da una società privata: v. sul punto per maggiori dettagli M. Gialuz, S. Quattrocolo, *AI and the Administration of Criminal Justice in Italy*, in *eRevue Internationale de Droit Pénal* 2023, 2, 1. V. altresì L. Algeri, *op. cit.*, 731; F. Basile, *Intelligenza artificiale e diritto penale*, cit., 12.

[35] Sul tema v., tra gli altri, T.W. Hung, C.P. Yen, *On the Person-Based Predictive Policing of AI*, in *Ethics and Information Technology* 2021, 23, 165 ss.; A.G. Ferguson, *The Rise of Big Data Policing*, cit., 35 ss.; L. Sommerer, *op. cit.*, 30 s.; S. Egbert, M. Leese, *Criminal Futures. Predictive Policing and Everyday Police Work*, Oxon-New York 2021, 28 ss.

[36] Su tale teoria v. A.G. Ferguson, *The Rise of Big Data Policing*, cit., 35 ss.

[37] V. *infra* 3.

[38] Riferimenti in A.G. Ferguson, *Policing Predictive Policing*, cit., 1140 s.; I. Mugari, E.E. Obioha, *op. cit.*, 6.

un *social network* comprensivo di tre *layers of offenders*: 1) *initial target offenders*; 2) *target offenders' associates*; 3) *associates of the target offenders' associates*. Si passa così all'ultima fase di carattere "operativo": la polizia contatta i soggetti inclusi nella rete sociale e procede a informare costoro circa il fatto di essere stati identificati «as a cause of violence in the city»; prospetta loro la possibilità di iniziare un percorso di "reinserimento sociale"[39]; al contempo, diffida i potenziali criminali dal violare legge, con importanti riflessi sanzionatori laddove essi siano poi ritenuti responsabili di un qualunque reato[40].

Un secondo sistema di *predictive policing* su cui pare utile soffermare l'attenzione è la *Strategic Subject List* (SSL), meglio nota come *Heat List* e successivamente denominata *Crime and Victimization Risk Model*, pure in questo caso basata sulla *social network analysis*[41].

Esso è stato adottato dalla polizia di Chicago (Illinois) per identificare i soggetti maggiormente a rischio di essere coinvolti, sia come autore sia come vittima[42], in episodi di violenza con armi da fuoco connessi alle *gang* o in un omicidio[43]. Il *software* esamina gli individui con precedenti penali e, in base al punteggio di pericolosità loro

[39] Questo può articolarsi nella partecipazione a programmi di *education*, *job training* ovvero in attività di sostegno nel caso in cui i soggetti in questione siano affetti da dipendenze. V. A.G. Ferguson, *Policing Predictive Policing*, cit., 1141, nota n. 197.

[40] È esemplificativo il caso, registratosi nel Western District del Missouri, di un soggetto sorpreso con un *bullet* nelle proprie tasche e perciò condannato a una pena detentiva della durata di ben quindici anni. V. sul punto V. A.G. Ferguson, *Policing Predictive Policing*, cit., 1141, nota n. 198; nonché più nel dettaglio Id., *Predictive Prosecution*, in *Wake Forest Law Review* 2016, vol. 51, no. 3, 719. È interessante notare che, ai sensi della legislazione in vigore nello Stato del Missouri, l'ipotesi contestata nel caso appena riportato ricade nel *class D felony* di *unlawful possession of a firearm*, il quale punisce un soggetto già resosi responsabile di un delitto (*felon*) con una pena fino a sette anni di reclusione e $10.000; tuttavia, qualora l'agente sia già stato condannato per un *dangerous felony*, il fatto integrerà un *class C felony*, punibile con la pena detentiva da tre a dieci anni e con la medesima pena pecuniaria prima indicata (cfr. https://revisor.mo.gov/main/OneSection.aspx?section=571.070).

[41] V. S. Egbert, M. Leese, *op. cit.*, 29, secondo cui la *Heat List* rappresenta il più importante esempio di analisi delle reti sociali. Su tale sistema v. inoltre L. Sommerer, *op. cit.*, 66 ss.; A.G. Ferguson, *Policing Predictive Policing*, cit., 1138 ss.; I. Mugari, E.E. Obioha, *op. cit.*, 6; A. Tucek, *Constraining Big Brother: The Legal Deficiencies Surrounding Chicago's Use of the Strategic Subject List*, in *The University of Chicago Legal Forum* 2018, 431 ss. che ne illustra nel dettaglio il funzionamento.

[42] Si esprime criticamente al riguardo L. Sommerer, *op. cit.*, 68, la quale sottolinea: «no distinction is made between potential perpetrators and potential victims, instead, they are pooled and provided with one risk score. It is therefore impossible for the police to use the score to understand whether they are dealing with a high-risk perpetrator or victim. As such, in case of doubt police officers will likely assume they are dealing with perpetrators to err on the side of caution».

[43] La scelta deriva dal numero elevato di questi episodi. V. il *report* di A.A. Braga, D.W. Webster, M.D. White, H. Saizow, *SMART Approaches to Reducing Gun Violence. Smart Policing Initiative Spotlight on Evidence-Based Strategies and Impacts*, marzo 2014.

attribuito[44], procede alla classificazione in diverse categorie. I dati considerati[45] attengono al numero di precedenti arresti (in particolare, per *violent offenses*, affiliazioni alle *gang* reati in materia di stupefacenti *e unlawful use of a weapon*), nonché alla circostanza di essere stati vittime in *shootings* o *aggravated battery or assault*[46], all'età del soggetto all'epoca dell'ultimo arresto e, infine, all'intensità della sua attività criminale nel tempo[47].

Anche qui, il processo si conclude con l'intervento, secondo le modalità sopra descritte, delle forze dell'ordine, che provvedono alla notifica della *custom notification letter*[48]. La procedura, dettagliata nello *Special Order S10-05* del dipartimento di polizia di Chicago, prevede – per quanto qui rileva – che, quando una persona identificata nella *Heat List* è nuovamente arrestata, la polizia raccomanderà ai pubblici ministeri di elevare le contestazioni più severe[49].

Concludiamo l'analisi soffermandoci sui *suspect-based systems* che, pur essendo stati classificati come sottocategoria dei sistemi *person-based*[50], si differenziano da questi ultimi poiché non sono basati su un *risk assessment* individuale. I sistemi in questione, analizzando i dati relativi a precedenti reati di carattere seriale (come, ad esempio, le rapine), tentano di cogliere alcuni tratti comuni a elaborare il profilo – anonimo – del possibile autore. Una simile operazione può avere una duplice utilità: in primo luogo, è funzionale a supportare le indagini relative a reati già consumati; inoltre, può avere una valenza "predittiva" rispetto alla futura commissione di reati da parte dello stesso soggetto.

---

[44] Il *predictive threat score* varia da zero a cinquecento.

[45] V. nello specifico Chicago Police Department, *Special Order S10-05*, 6.10.2015 che disciplina l'intera procedura. In dottrina, v. A. Tucek, *op. cit.*, 432.

[46] Si allude sostanzialmente a reati di minacce e percosse aggravate.

[47] V. Chicago Police Department, *Special Order S10-05*, cit., che indica, letteralmente, "*the degree to which his criminal activities are on the rise*"; tuttavia, il provvedimento nulla aggiunge circa le relative modalità di calcolo. Questo fattore sembrerebbe essere funzionale a ponderare il peso assegnato alla precedente attività criminale, nel senso che laddove, ad esempio, gli arresti riportati da un soggetto siano tutti recenti, gli sarà attribuito uno *score* più alto rispetto a colui che vanta lo stesso numero di precedenti ma più risalenti, in quanto il suo comportamento criminale "*on the rise*" denoterà la sua maggiore pericolosità. In termini analoghi, v. B. Posadas, *How strategic is Chicago's "Strategic Subjects List"? Upturn investigates*, in *Medium*, 22.6.2017.

[48] La lettera contiene l'indicazione dei precedenti contatti del soggetto con la "giustizia penale" (quindi, non solo dei suoi precedenti), offre la possibilità di usufruire dei servizi sociali, ma contiene altresì la prospettazione delle conseguenze cui costui andrà incontro laddove non si astenesse dal delinquere. Sul punto, v. A. Tucek, *op. cit.*, 435 s.; A.G. Ferguson, *Predictive Prosecution*, cit., 715.

[49] V. Chicago Police Department, *Special Order S10-05*, cit.; A.G. Ferguson, *Predictive Prosecution*, cit., 717.

[50] V. *supra* § 2.

Sebbene i *software* in questione non siano molto diffusi negli Stati Uniti, un esempio promettente è stato ideato da un esponente della Questura di Milano e poi ulteriormente sviluppato grazie all'impiego dell'intelligenza artificiale e del *machine learning*. Alludiamo a *KeyCrime*[51], un sistema che esamina tutte le informazioni, anche le più dettagliate, ricavabili dai reati di natura seriale già realizzati – circostanze di luogo e di tempo, abbigliamento e gestualità del colpevole, testimonianze dei presenti, etc. – per delineare il "profilo" del relativo autore e tentare di predire i suoi successivi comportamenti.

Inquadrati i diversi sistemi ed esaminate le modalità di funzionamento, è giunto il momento di concentrarsi sui profili problematici emersi nella prassi.

3. L'esperienza applicativa in materia di polizia predittiva si è caratterizzata per una prima fase di diffusa applicazione di questi sistemi nei dipartimenti di polizia nordamericani[52]. Tuttavia, ciò è avvenuto in assenza di una regolamentazione sul loro utilizzo e – prima ancora – di un confronto preliminare con giuristi, *policy makers*, etc.[53]; ciò ha ben presto sollevato problematiche di vario genere che, in diversi casi, hanno portato alla dismissione dei *software*.

---

[51] V. L. Algeri, *op. cit.*, 731; F. Basile, *Intelligenza artificiale e diritto penale*, cit., 12; A. Bonfanti, *op. cit.*, 3; G. Padua, *Intelligenza artificiale e giudizio penale: scenari, limiti e prospettive*, in *PPG* 2021, 1492; C. Parodi, V. Sellaroli, *op. cit.*, 56 s.; nonché, da ultimo, M. Gialuz, S. Quattrocolo, *op. cit.*, 4. *KeyCrime*, oggi di proprietà di un'azienda privata, si è evoluto nel *Dynamic Evolving Learning Integrated Algorithm* (Delia). Inoltre, proprio sulla base di tale sistema, il Dipartimento di Pubblica Sicurezza del Ministero dell'Interno aveva lanciato, nel 2023, il progetto *Giove*, un sistema di elaborazione e analisi automatizzata, che ricalca le modalità di funzionamento del *tool* in questione e sul quale era stata presentata una interrogazione parlamentare (v. sul punto E. Pietrocarlo, *op. cit.*, 120, nota n. 49).

[52] V. per tutti E. Joh, *Ethical AI in American Policing*, in *Notre Dame Journal on Emerging Technologies* 2022, vol. 3, no. 2, 10.

[53] In questo senso, ancorché rispetto all'esperienza della Germania, v. L. Sommerer, *op. cit.*, 35, la quale osserva problematicamente come *PreCobs*, il sistema *place-based* di cui si è detto (*supra* § 2), sia stato sviluppato e implementato senza coinvolgimento della comunità scientifica, che avrebbe invece consentito di affrontare e risolvere sin da subito alcune criticità. Nello stesso senso, v. A.G. Ferguson, *Policing Predictive Policing*, cit., 1143. Un primo tentativo di regolamentazione della *predictive policing* si è avuto grazie alla campagna *Community Control Over Police Surveillance* (CCOPS) lanciata dall'ACLU – acronimo di *American Civil Liberties Union* – nel settembre 2016. In sostanza, la stessa promuoveva l'adozione, da parte dei governi locali, di ordinanze recanti la disciplina di approvazione preventiva di ogni tecnologia di sorveglianza, offrendo altresì un modello ideale di provvedimento (c.d. *CCOPS bill*). L'iniziativa, tuttavia, ha avuto un seguito "a macchia di leopardo", oltre al fatto che le ordinanze concretamente adottate non hanno riprodotto fedelmente il modello elaborato dall'ACLU, compromettendo talvolta l'effettività delle stesse. Per un'analisi comparativa delle ordinanze in questione, v. il *white paper* di T. Chivukula, A.Takemoto, *Local Surveillance Oversight Ordinances*, Samuelson Law, Technology & Public Policy Clinic – Berkeley School of Law, University of California, febbraio 2021 (https://www.law.berkeley.edu/case-project/local-surveillance-oversight-ordinances-white-paper/) nonché, volendo, anche per ulteriori riferimenti, E. Pietrocarlo, *op. cit.*, 127 ss. Nell'ottobre 2022 l'*Office of Science and*

Alcuni profili di criticità sono comuni a tutte le tipologie di *tools* e saranno, pertanto, oggetto di una trattazione unitaria. Riserveremo invece un'analisi separata alle questioni sollevate solo da taluni degli strumenti esaminati.

Un primo, comune ordine di problemi attiene ai *dati* di cui si alimenta l'algoritmo, non solo avuto riguardo ai possibili errori nelle fasi di raccolta, selezione e inserimento nel sistema[54] – i quali si rifletteranno inevitabilmente sulla validità del risultato – ma, soprattutto, in relazione alla loro *qualità*[55]. Sul punto la preoccupazione è duplice: da un lato, il dato potrebbe non essere effettivamente rappresentativo della realtà e, dunque, la predizione potrebbe risultare falsata (basti pensare alla "cifra nera" esistente rispetto a *violent crime* come la violenza domestica)[56]; dall'altro lato, molti dei dati processati dal *software* (come i contatti avuti con la polizia) sono "prodotti" dagli stessi agenti, con il rischio – prospettato in dottrina[57] – che possano essere viziati da *implicit bias* rispetto a talune minoranze.

Technology Policy (OSTP) presso la Casa Bianca ha pubblicato il *"Blueprint for an AI Bill of Rights. Making Automated Systems Work for the American People"*, un libro bianco che enuncia cinque principi per il corretto utilizzo dei sistemi automatizzati di AI che dovrebbero guidare il futuro sviluppo di politiche in materia a salvaguardia dei diritti civili. Per la relativa analisi, sia consentito rinviare ancora a E. Pietrocarlo, *op. cit.*, 130 ss. Da ultimo, il 30 ottobre 2023, il Presidente Biden ha adottato l'*Executive Order 14110 "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"*, il primo atto volto a regolare organicamente l'AI, il quale vincola (soltanto) le entità federali cui è rivolto (e non la generalità dei consociati). Oltre a una serie di definizioni rilevanti in materia, l'EO enuncia le otto finalità cui mira il sistema statunitense (quali, la sicurezza, l'innovazione responsabile; la promozione e la protezione del lavoro negli Stati Uniti; la tutela dell'uguaglianza e dei diritti umani; la protezione dell'uso quotidiano dell'AI; la protezione dei dati personali; il reclutamento nelle agenzie federali di personale qualificato; l'assunzione di un ruolo guida nel mondo), le quali si dovranno tradurre nello svolgimento di precise politiche pubbliche, che i diversi dipartimenti governativi dovranno attuare mediante specifiche azioni ed entro determinate scadenze. Con particolare riguardo al tema di nostro interesse, si prevede che le politiche volte ad assicurare l'uguaglianza e i diritti umani siano di competenza dell'Attorney General, il quale, entro un anno dall'entrata in vigore del provvedimento presidenziale, dovrà, tra l'altro, consegnare una relazione al Presidente medesimo contenente dati sull'uso dei sistemi di AI in diversi ambiti, tra cui sono incluse le attività di polizia predittiva. Per un primo commento, v. E. Carloni, *Il sentiero si fa camminando: la strategia statunitense per intelligenze artificiali sicure ed affidabili*, in *GiornDirAmm* 2024, 1, 135 ss.

[54] A.G. Ferguson, *Policing Predictive Policing*, cit., 1145.

[55] V. M.G. Peluso, *Intelligenza Artificiale e dati di qualità: la tecnologia come valido alleato*, in *MediaLaws* 2022, 2, 326 ss.

[56] V. sul punto A.G. Ferguson, *Policing Predictive Policing*, cit., 1147 che riporta alcuni dati del Department of Justice.

[57] V., tra i molti, P.J. Brantingham, *The Logic of Data Bias and its Impact on Place-Based Predictive Policing*, in *Ohio State Journal of Criminal Law* 2018, vol. 15, 475 ss; A.G. Ferguson, *Policing Predictive Policing*, cit., 1148 s.; S. Egbert, M. Leese, *op. cit.*, 190 ss.; S. Barocas, A. Selbst, *Big Data's Disparate Impact*, in *California Law Review* 2016, vol. 104, 673 ss.; F. Mirò-Llinares, *Predictive policing: Utopia or dystopia? On attitudes towards the use of big data algorithms for law enforcement*, in *IDP: Revista de Internet, derecho y política* 2020, n. 30, 9; nonché, in particolare, S. Mayson, *Bias In, Bias Out*, in *The Yale Law Journal* 2019, vol. 128, 2218 ss.; C. Burchard, *L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*, in *RIDPP* 2019, 1932 s. Con

Un ulteriore tema – che sicuramente si pone laddove gli *input* siano costituiti da (enormi quantitativi) di dati personali, come ad esempio nei sistemi *person-based*[58] – è quello relativo all'esigenza di regolamentare i processi di raccolta, analisi, conservazione e cancellazione degli stessi [59] e, al contempo, evitare possibili degenerazioni che diano vita a forme di sorveglianza di massa.

Un secondo ordine di problemi concerne l'opacità dell'algoritmo, ovverosia la non ricostruibilità *ex post* del percorso compiuto per addivenire a un determinato risultato, sia a causa di ragioni tecniche – si tratta del noto *black box problem* dovuto al fatto che gli algoritmi basati sull'intelligenza artificiale si distaccano dalle istruzioni iniziali per apprendere autonomamente – sia di ragioni economiche, dal momento che le aziende

particolare riguardo all'esperienza pratica di Chicago, v. l'*audit* (J.S. Hollywood, K.N. McKay, D. Woods, D. Agniel, *Real-Time Crime Centers in Chicago. Evaluation of the Chicago Police Department's Strategic Decision Support Centers*, Santa Monica 2019 (https://www.rand.org/pubs/research_reports/RR3242.html) eseguita su tutte le tecnologie basate sui *big data* utilizzate da tale dipartimento di polizia – ivi inclusa la c.d. *Heat List* – all'esito della quale è stato rilevato, *inter alia*, come i dati di cui siffatto sistema si alimentava – in particolare, quelli relativi ai precedenti contatti con la polizia (es. arresti) – rischiavano di essere agevolmente influenzati da *bias*, trattandosi di elementi ottenuti dai rapporti redatti al tempo dagli stessi agenti. Peraltro, è proprio a seguito della pubblicazione di tale documento che, nel novembre 2019, la "lista calda" è stata abbandonata: v. S. Charles, *CPD decommissions "Strategic Subject List"*, in *Chicago Sun Times*, 27.1.2020 (https://chicago.suntimes.com/cityhall/2020/1/27/21084030/chicago-police-strategic-subject-list-party-to-violence-inspector-general-joeferguson). Sul punto in dottrina, v. A.G. Ferguson, *Surveillance and the Tyrant Test*, in *The Georgetown Law Journal* 2021, vol. 110, no. 2, 228 s. La potenzialità discriminatoria della polizia predittiva è stata altresì denunciata nella giurisprudenza statunitense: v. la *concurring opinion* del giudice Thacker espressa nella decisione, del 15.7.2020, della Corte d'appello del Fourth Circuit relativa al caso *United States of America v. Billy Curry, Jr.*, secondo cui la *predictive policing* «is no longer the shiny new object it may once have appeared to be, but instead has revealed itself to be tarnished with racial bias». Per il testo della pronuncia v. https://www.ca4.uscourts.gov/opinions/184233A.P.pdf, mentre per un commento, v. S. Lonati, Predictive policing*: dal disincanto all'urgenza di un ripensamento*, in *MediaLaws* 2022, 2, 308 s.

[58] Sulla possibilità di qualificare come "personali" i dati cui si alimentano i sistemi *place-based*, v. O. Lynskey, *Criminal Justice Profiling and EU Data Protection Law: Precarious Protection from Predictive Policing*, in *International Journal of Law in Context* 2019, vol. 15, 168 ss. e, in particolare, 171 nonché diffusamente *infra* § 4.

[59] V. A.G. Ferguson, *Policing Predictive Policing*, cit., 1185. Tra gli italiani, v. G. Ubertis, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *DPenCont* 2020, 4, 81 s.; F. Basile, *Intelligenza artificiale e diritto penale*, cit., 13. Per ulteriori spunti di riflessione sul tema, v. E. Carpanelli, *Il ricorso all'intelligenza artificiale nel contesto di attività di* law enforcement *e di operazioni militari: brevi riflessioni nella prospettiva del diritto internazionale*, in *www.dirittopenalecontemporaneo.it,* 2022, 1, 386. Un tema connesso alla *data protection* è quello affrontato dalla sentenza del Tribunale costituzionale federale tedesco del 16.2.2023 che ha dichiarato incostituzionali due disposizioni di leggi regionali di polizia che consentivano l'analisi automatizzata di grandi *database* per la prevenzione di reati, in quanto lesive del diritto all'autodeterminazione informativa, ovverosia il «diritto del cittadino di disporre liberamente della diffusione e dell'uso dei dati che lo riguardano»: così R. Orlandi, *Uso poliziesco dell'intelligenza artificiale. L'insegnamento del Bundesverfassungsgericht*, in *CP* 2023, 7-8, 2170, cui si rinvia per una compiuta analisi della pronuncia; riferimenti alla pronuncia in discorso anche in J. Sprenger, D. Brodowski, *op. cit.*, 144 ss.

proprietarie dei *software* oppongono il segreto industriale alle richieste di chiarificazioni sul punto[60].

Inoltre, il difetto di trasparenza si riflette sul piano dell'*accountability*[61], nel senso che i "bersagli" delle predizioni algoritmiche non sono messi nelle condizioni di contestare, ai soggetti responsabili (o, più correttamente, *accountable*)[62], gli effetti lesivi che vengono così a prodursi a loro danno.

Anche sul diverso piano delle modalità di funzionamento degli algoritmi sorgono profili di criticità, potendo presentare una potenzialità discriminatoria laddove non implementati secondo schemi oggettivi e scevri da pregiudizi[63]. Peraltro, ciò assume particolare rilievo rispetto al *machine learning* poiché la capacità di autoapprendimento è suscettibile, come noto, di amplificare ulteriormente i *bias* di fondo.

Infine, si osserva come le diverse problematiche sin qui descritte finiscano per compromettere l'affidabilità e l'accuratezza delle predizioni; il che desta in particolare allarme avuto riguardo ai *person-based systems* allorché si qualifichino "pericolosi" soggetti che in realtà non sono da considerare tali[64], tenuto altresì conto delle conseguenze che abbiamo visto derivare in quei sistemi da tale classificazione. Peraltro, non si prevede neppure un controllo umano, che possa in qualche modo rimediare a eventuali errori del *software*, oltre che rappresentare una garanzia per l'individuo.

---

[60] V. tra gli altri E. Joh, *Ethical AI in American Policing*, cit., 11 s.; T.W. Hung, C.P. Yen, *op. cit.*, 166 s. V. altresì P. Severino, *Le implicazioni dell'intelligenza artificiale nel campo del diritto con particolare riferimento al diritto penale*, in *Intelligenza artificiale. Politica, economia, diritto, tecnologia*, a cura di P. Severino, Roma 2022, 97. Sul tema della trasparenza v., nella dottrina italiana, F. Consulich, *"Flash offenders". Le prospettive di "accountability" penale nel contrasto alle intelligenze artificiali devianti*, in *RIDPP* 2022, 3, 1027.

[61] V. per tutti A.G. Ferguson, *Policing Predictive Policing*, cit., 1169.

[62] Sulla difficile individuazione del significato dell'*accountability* e sulla differenza rispetto al concetto di responsabilità, v. A. Giannini, *Intelligenza artificiale,* human oversight *e responsabilità penale: prove d'impatto a livello europeo*, in *Discrimen*, 21.11.2021, 7 ss.

[63] E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, in *Washington Law Review* 2014, vol. 89, 58; I. Mugari, E.E. Obioha, *op. cit.*, 10.

[64] Al riguardo, v. l'*audit* dell'*Inspector General* (v. https://www.lapdpolicecom.lacity.org/031219/BPC_19-0072.pdf) sull'utilizzo del sistema di polizia predittiva LASER implementato a Los Angeles, da cui è emersa, tra l'altro, proprio l'inaffidabilità delle predizioni, in quanto nel calcolo del *risk score* spesso non si teneva conto di una reale corrispondenza tra tasso di pericolosità e precedenti penali del soggetto in questione. Invero è stato rilevato che il 44% dei soggetti inclusi nella "lista calda" (in questo caso si trattava del c.d. *Chronic Offenders Bulletin*) era stato arrestato solo una volta o, addirittura, mai; ciò poiché il sistema in esame considera, ai fini dell'inserimento nell'elenco, anche le vittime di alcuni reati. Per ulteriori riferimenti a LASER, volendo, v. E. Pietrocarlo, *op. cit.*, 122.

Al di là di tali profili comuni, con particolare riguardo ai *place-based systems* è stato sottolineato il rischio di un c.d. *feedback loop*[65], causato dalle stesse modalità di funzionamento dei *software*. In particolare, in virtù della concentrazione dei controlli in una determinata area urbana originariamente qualificata "a rischio" dal *tool*, la polizia rileverà ragionevolmente un determinato numero di reati; reinserendo poi nel sistema i dati relativi a tali ultime attività criminali, si finirà per aumentare il tasso di rischio di quella zona, con il risultato di indirizzare lì tutte le risorse – trascurando, di conseguenza, altri quartieri – e di sottoporre a una sorveglianza costante i relativi residenti.

Ulteriori, rilevanti criticità riguardano poi, nello specifico, i *person-based systems*.

A non convincere sono anzitutto i parametri considerati per selezionare i soggetti a rischio, ai fini del loro inserimento nelle "liste calde": sono infatti inclusi anche coloro che hanno legami soltanto indiretti con individui coinvolti, anche come vittime, in episodi di violenza[66]. In realtà, ci sembra necessario un quadro indiziario più robusto affinché un soggetto possa essere qualificato come "a rischio", oltre alla previsione di meccanismi che consentano al *predicted offenders* di contestare il suo avvenuto inserimento negli elenchi in questione – dei quali non vi è invece traccia nella prassi.

A conferma della problematicità dei sistemi in discorso, si consideri altresì la potenziale influenza che la documentazione predisposta dalla polizia all'esito della discutibile analisi algoritmica può esercitare sulle determinazioni della pubblica accusa[67]. Invero, nell'ambito della *Heat List* di Chicago, si prevede il coinvolgimento diretto dei *prosecutors*, i quali potranno non solo partecipare agli incontri ma anche tenere conto della *custom notification letter*, elaborata sulla base di un'analisi algoritmica che, come detto poc'anzi, presenta diversi profili di criticità (qualità dei dati, *bias*, accuratezza, etc.). Ci paiono evidenti da questo angolo visuale i profili di frizione con la presunzione di innocenza.

---

[65] Una simile criticità è plasticamente illustrata nel *Blueprint on AI* della Casa Bianca allorché viene enunciato il principio per cui i sistemi di AI devono essere *"Safe and Effective"*. Nella dottrina italiana, v. in particolare P. Severino, *Intelligenza artificiale e diritto penale*, cit., 541 s.; F. Basile, *Intelligenza artificiale e diritto penale*, cit., 30.

[66] V. A.G. Ferguson, *Predictive Prosecution*, cit., 720 e A. Tucek, *op. cit.*, 427 s. che riportano, in termini critici, il caso di un giovane con alcuni dei precedenti penali, che tuttavia consistevano in una sola condanna per *misdemeanor* e alcuni arresti minori, e amico di un soggetto *vittima* di un conflitto a fuoco, che per ciò solo è stato inserito nella c.d. *Heat List* di Chicago e poi è risultato destinatario dei vari avvertimenti sulle possibili conseguenze di un suo eventuale agire illecito, compresa l'inflizione di una pena più severa.

[67] È critico al riguardo A.G. Ferguson, *Predictive Prosecution*, cit., 719.

In ultima analisi, sono gli stessi meccanismi di diffida a destare perplessità. Abbiamo visto come essa si sostanzi nell'invito ad astenersi dal commettere ulteriori reati, pena la possibilità di incorrere in conseguenze di rilievo, tra cui l'irrogazione di una sanzione più severa in ragione del precedente avvertimento – e, in effetti, la prassi testimonia come tale eventualità si sia poi concretizzata dando luogo a un trattamento punitivo sproporzionato alla gravità del fatto[68].

È tuttavia doveroso sottolineare come un simile procedimento – ovverosia l'esecuzione di una diffida, a fronte di una previa valutazione di pericolosità individuale da parte dell'autorità di pubblica sicurezza, unitamente alla previsione di un trattamento sanzionatorio deteriore – possa, a certe rigorose condizioni, essere legittimo, come del resto dimostrano alcuni istituti presenti nell'ordinamento italiano[69]. Ciò che dalla nostra prospettiva rende insuperabili le criticità riscontrate rispetto ai sistemi *person-based* americani è, da un lato, l'assenza di una precisa regolamentazione e di garanzie nei confronti del destinatario nonché, dall'altro, il rischio di violare i principi fondamentali della materia penale (quali, tra gli altri, la proporzione del trattamento sanzionatorio e la presunzione d'innocenza). Al contrario, come peraltro suggerito da alcuni autori[70], le altre problematiche emerse e,

[68] V. nota n. 40.

[69] Il riferimento è anzitutto all'avviso orale di cui all'art. 3 del c.d. Codice Antimafia, misura di prevenzione applicata dal questore e consistente in una diffida a non violare la legge nei confronti di un soggetto che – oltre a rientrare in una delle ipotesi di pericolosità generica di cui all'art. 1 del Codice – sia pericoloso per la pubblica sicurezza. Per ciò che qui rileva, laddove il sottoposto a una misura di prevenzione personale – e, dunque, anche il destinatario dell'avviso – commetta, ad esempio, uno dei reati indicati dall'art. 71 del medesimo Codice, verrà in rilievo una circostanza aggravante, la fattispecie sarà procedibile d'ufficio, si potrà procedere all'arresto fuori dei casi di flagranza (in relazione ai delitti, tra quelli richiamati, per cui è ammesso l'arresto in flagranza) e si applicherà una misura di sicurezza detentiva in aggiunta alla pena prevista per il reato. Ulteriori conseguenze sanzionatorie negative sono previste agli artt. 72 e 73, allorché, rispettivamente, il destinatario della misura commetta reati concernenti le armi e gli esplosivi di cui agli artt. 1 e 2, co. 1 e 2, della l. 18.4.1975, n. 110, ovvero specifiche violazioni del Codice della strada. Su tale misura, v. A. Chelo, sub *art. 3, d. lgs. n. 159/2011*, in *Commentario breve al Codice Antimafia e alle altre procedure di prevenzione*, a cura di G. Spangher, A. Marandola, Milano 2019, 24 ss.; F. Basile, *Manuale delle misure di prevenzione. Profili sostanziali*, Torino 2020, 82 ss. Un meccanismo analogo è previsto nell'ambito dell'ammonimento del questore in materia di atti persecutori di cui all'art. 8 del d.l. 23.2.2009, n. 11 (convertito con modificazioni dalla l. 23.4.2009, n. 38), in quanto qualora l'ammonito commetta fatti di *stalking*, incorrerà in un aggravamento di pena e muterà il regime di procedibilità (da querela a officioso). Sul tema v. R. Bricchetti, L. Pistorelli, *Istanza di ammonimento: una prima forma di tutela*, in *GD* 2009, 10, 69 ss.; A.M. Gasparre, *L'istituto giuridico dell'ammonimento del questore per l'appartenente alla Polizia di Stato: peculiarità e conseguenze*, in *RP* 2021, 4, 309 ss. Per un raffronto tra questi istituti e i sistemi di polizia predittiva americani, sia consentito rinviare a E. Pietrocarlo, *op. cit.*, 133 ss.

[70] V. in particolare le proposte di M. Bagaric, J. Svilar, M. Bull, D. Hunter, N. Stobbs, *The Solution to the Pervasive Bias and Discrimination in the Criminal Justice System: Transparent and Fair Artificial Intelligence*, in *American Criminal Law Review* 2022, vol. 59, 125 ss. Per una generale ricostruzione dei "correttivi" proposti nella letteratura americana, sia consentito il rinvio a E. Pietrocarlo, *op. cit.*, 126 s.

per così dire, "connaturate" agli algoritmi e alle loro modalità di funzionamento possono essere neutralizzate o, quantomeno, minimizzate attraverso l'imposizione di specifiche prescrizioni. E – come subito vedremo – è proprio questa l'impostazione seguita dall'ultima versione del Regolamento europeo.

La ricostruzione sin qui compiuta ci permetterà ora di meglio comprendere il cammino dell'*AI Act*, con specifico riferimento al tema qui indagato, nonché di chiarire quale sia l'ambito riservato ai sistemi di *predictive policing*[71].

4. La prima bozza della proposta di Regolamento europeo sull'intelligenza artificiale è stata elaborata dalla Commissione il 21 aprile 2021[72]. L'ambito in cui era possibile riscontrare un riferimento alle applicazioni di polizia predittiva era quello dei sistemi ad alto rischio di cui all'Allegato III, paragrafo 6, riservato alle «Attività di contrasto»[73]. A venire in rilievo era, in particolare, il disposto delle lettere *a*, *e*, *f* e *g*, che esamineremo subito per comprendere a quali concrete tipologie di strumenti si alludeva.

La lettera *a* si riferiva ai «sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per effettuare valutazioni individuali dei rischi delle persone fisiche al fine di determinare il rischio di reato o recidiva in relazione a una persona fisica o il rischio per vittime potenziali di reati»; categoria nella quale potevano certamente essere ricondotti i sistemi *person-based* che comportano un *risk assessment* individuale.

Quanto alla lettera *e*, essa includeva i sistemi che, pur non implicando una valutazione di pericolosità di un soggetto, sono utilizzati dalle autorità di *law enforcement* «per prevedere il verificarsi o il ripetersi di un reato effettivo o potenziale sulla base della profilazione delle persone fisiche di cui all'articolo 3, paragrafo 4, della direttiva (UE) 2016/680 o per valutare i tratti e le caratteristiche della personalità o il comportamento criminale pregresso di persone fisiche o gruppi».

Per verificare se in tale insieme potevano essere inquadrati i *place-based systems*, occorre anzitutto chiarire il significato di "profilazione" a mente della *Law Enforcement Directive* (espressamente richiamata dalla proposta della Commissione). Essa è

[71] V. rispettivamente *infra* §§ 4 e 5.

[72] Commissione Europea, *Proposta per un Regolamento che stabilisce regole armonizzate sull'intelligenza artificiale, COM(2021) 206 final*, 21.4.2021 (https://digital-strategy.ec.europa.eu/en/library/proposalregulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence).

[73] Nella versione in lingua inglese "*Law enforcement*". Per un commento alla iniziale proposta della Commissione dalla prospettiva penalistica, v. A. Lavorgna, G. Suffia, *La nuova proposta europea per regolamentare i Sistemi di Intelligenza Artificiale e la sua rilevanza nell'ambito della giustizia penale*, in *DPenCont* 2021, 2, 89 ss.

definita come «qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica»[74]. La profilazione evidentemente poggia sul concetto di "dati personali". Non resta allora che confrontarsi con la relativa definizione di cui all'art. 3, paragrafo 1, della medesima Direttiva, ai sensi del quale per "dato personale" si intende: «qualsiasi informazione riguardante una persona fisica identificata o identificabile, (l'«interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, in particolare con riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici dell'identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale di tale persona fisica»[75].

È dunque dalla qualificazione come personali o meno dei dati su cui si basa il processo di elaborazione dei *place-based systems* che sembra dipendere la loro riconducibilità alla categoria di sistemi ad alto rischio delineata dalla lettera *e*.

In senso affermativo, taluno ha osservato[76] come, sulla base di alcune pronunce della Corte di Lussemburgo e delle indicazioni ricavabili da fonti di *soft law*, si potrebbe far leva su un'interpretazione particolarmente estesa del dato personale che vada oltre il mero "contenuto" del dato medesimo – cioè la sua attinenza a una determinata persona – e guardi piuttosto all'impatto del suo utilizzo su un individuo[77]. Di conseguenza, sarebbe possibile sostenere che anche i dati di cui si alimentano i

[74] Così l'art. 3, paragrafo 4, dir. 2016/680/UE (c.d. Direttiva *Law Enforcement*).

[75] Art. 3, paragrafo 1, dir. 2016/680/UE; v. inoltre l'art. 4, n. 1) del GDPR che contiene il medesimo disposto. Sul concetto di "dato personale", in dottrina v. F. Guerrieri, sub *art. 4 GDPR*, in *GDPR e Normativa Privacy Commentario*, a cura di G.M. Riccio, G. Scorza, E. Belisario, Milano 2022, 42 ss.

[76] V., anche per gli opportuni riferimenti, O. Lynskey, *op. cit.*, 168 ss. e 171, che ha prospettato questa possibilità nell'ambito di un'analisi volta a verificare l'applicabilità della Direttiva *Law Enforcement* alle diverse tipologie di sistemi di polizia predittiva.

[77] V. Gruppo di Lavoro Articolo 29 per la Protezione dei Dati Personali, *Parere 4/2007 sul concetto di dati personali*, p. 10, ove si afferma che «per stabilire se i dati "concernono" una persona, dovrebbe ricorrere un elemento di "contenuto" OPPURE di "finalità" OPPURE di "risultato"». Con riguardo a tale ultimo elemento, si specifica (p. 11) che «è possibile considerare che i dati "concernono" una persona quando il loro impiego può avere un impatto sui diritti e sugli interessi di quella persona, tenendo conto di tutte le circostanze del caso di specie. Si noti che non è necessario che il risultato potenziale abbia un impatto importante. È sufficiente che la persona sia trattata in modo diverso rispetto ad altre in seguito al trattamento di tali dati».

place-based systems sono "personali" in ragione della loro potenziale incidenza su coloro che si trovano negli *hot spot* identificati dal *software*. Da qui la conclusione è immediata: trattandosi di strumenti che effettuano un trattamento automatizzato di dati personali – ovverosia una profilazione – per prevedere il verificarsi o il ripetersi di un reato effettivo o potenziale, essi sarebbero rientrati tra i sistemi ad alto rischio di cui alla lettera *e*.

Tuttavia, potrebbe apparire – come è stato rilevato – controintuitivo[78] affermare che i dati di cui si alimentano *software* del genere di *PredPol* (relativi a tipologia, data, orario e luogo di commissione di un reato) possano qualificarsi come "personali". Anche a voler ammettere che collegare, ad esempio, i dati relativi all'ubicazione di un crimine a un individuo sia ipoteticamente possibile e si tratti, dunque, di dati relativi a una persona "identificabile", si potrebbe obiettare che ciò non soddisfi il *"likely reasonable" standard*, il quale rappresenta il metro indicato dalla stessa Direttiva al Considerando n. 21 per stabilire l'identificabilità di una persona fisica[79].

In definitiva, ci sembra che la specifica tipologia di dati considerati dai sistemi *place-based* – ossia di carattere assolutamente "impersonale" – unitamente al fatto che gli stessi sono rivolti a un'analisi, per così dire, "situazionale" – cioè delle aree a rischio reato – la quale prescinde da valutazioni attinenti a individui debba condurre a negare la possibilità di qualificare i dati in questione come "personali" e, di riflesso, non si è in presenza di una profilazione. Ad ogni modo, la loro riconduzione alla lettera *e* pare potersi fondare sul riferimento a quei sistemi, richiamati nella seconda parte, che valutano il comportamento criminale pregresso.

Chiarite le diverse, possibili collocazioni dei *place-based systems*, possiamo passare all'esame della lettera *f*, secondo cui potevano considerarsi "ad alto rischio" «i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per la profilazione delle persone fisiche [...] nel corso dell'indagine, dell'accertamento e del perseguimento di reati». Infine, come categoria – potremmo dire – "residuale", la lettera *g* ricomprendeva «i sistemi di IA destinati a essere utilizzati per l'analisi criminale riguardo alle persone

[78] Così O. Lynskey, *op. cit.*, 171.

[79] V. ancora O. Lynskey, *op. cit.*, 171. Il Considerando n. 21 dir. 2016/680/UE a mente del quale: «[...] Per stabilire l'identificabilità di una persona fisica, è opportuno considerare tutti i mezzi, come l'individuazione, di cui il titolare del trattamento o un terzo può ragionevolmente avvalersi per identificare detta persona direttamente o indirettamente. Per accertare la ragionevole probabilità di utilizzo dei mezzi per identificare la persona fisica, si dovrebbe prendere in considerazione l'insieme dei fattori obiettivi, tra cui i costi e il tempo necessario per l'identificazione, tenendo conto sia delle tecnologie disponibili al momento del trattamento, sia degli sviluppi tecnologici [...]».

fisiche, che consentono alle autorità di contrasto di eseguire ricerche in set di dati complessi, correlati e non correlati, resi disponibili da fonti di dati diverse o in formati diversi, al fine di individuare modelli sconosciuti o scoprire relazioni nascoste nei dati».

Alla luce di tale articolato quadro normativo, ci pare fossero da includere nel perimetro dei sistemi ad alto rischio i *software suspect-based*, cioè quelli volti all'analisi delle informazioni ricavabili da una serialità criminale già realizzata al fine di tracciare il profilo del possibile autore e, eventualmente, di predire il suo comportamento futuro. La caratteristica di questi strumenti risiede infatti nel loro impiego non in una fase precedente alla commissione di un reato, bensì in quella prettamente investigativa a fronte della integrazione di condotte penalmente rilevanti[80]. In altre parole, la loro finalità non pare tanto quella di prevenire reati identificando i potenziali autori, quanto piuttosto quella di offrire un supporto alle tradizionali attività di indagine relative a crimini seriali e, segnatamente, all'analisi delle tracce lasciate dal relativo autore, la quale sarà senz'altro più rapida e completa in ragione delle straordinarie capacità analitiche dell'AI.

Questo era, insomma, il destino riservato alla polizia predittiva nella proposta della Commissione; tuttavia, l'eco dei problemi sorti nell'esperienza degli Stati Uniti ha condotto il Parlamento a presentare, nel giugno del 2023, diversi emendamenti nel senso di introdurre un divieto per la gran parte degli strumenti in esame, nel timore che potessero riproporsi gli "effetti collaterali" registratisi oltreoceano[81]. Vediamo nel dettaglio la portata delle modifiche in questione.

Innanzi tutto, è stata espunta dall'elenco dei sistemi ad alto rischio di cui all'Allegato III, paragrafo 6, la lettera *a* relativa agli strumenti di polizia predittiva *person-based* basati su una valutazione di pericolosità, nonché la lettera *e* riguardante i sistemi impiegati per prevedere il verificarsi o il ripetersi di un reato effettivo o potenziale sulla base della profilazione o per valutare i tratti e le caratteristiche della personalità o il comportamento criminale pregresso.

[80] Nell'esperienza della Questura di Milano, il *software* in esame veniva infatti (utilmente) impiegato a seguito della realizzazione di più reati seriali (in particolare, rapine presso farmacie). Grazie all'analisi da questo eseguita, era possibile delineare il profilo del relativo autore e identificare quelli che sarebbero stati i suoi più probabili obiettivi successivi, ivi allocando di conseguenza i controlli delle forze dell'ordine.

[81] Un simile scenario era indubbiamente prevedibile alla luce della posizione assunta dal Parlamento nella Risoluzione del 6.10.2021 di cui si è detto *supra* § 1, nota n. 2, che al paragrafo 24 riporta problematicamente l'esperienza degli Stati Uniti in materia di polizia predittiva opponendosi al suo utilizzo da parte degli Stati membri.

Di conseguenza, tali disposizioni sono state fatte confluire – con alcune modifiche – nella nuova lettera *d-bis* dell'art. 5 della proposta concernente le pratiche vietate. Una simile scelta si giustificava, secondo il Parlamento, in considerazione del fatto che la *predictive policing* violerebbe la dignità umana e la presunzione di innocenza nonché comporterebbe un accentuato rischio di discriminazione. Nello specifico, il divieto riguardava i sistemi volti a «effettuare valutazioni di rischio riguardo a persone fisiche o gruppi di persone fisiche al fine di valutare il rischio di reato o di recidiva di una persona fisica» – dunque i *person-based systems* – «o per prevedere il verificarsi o il ripetersi di un reato o di un illecito amministrativo effettivo o potenziale sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della personalità, compresa l'ubicazione della persona, o del comportamento criminale pregresso di persone fisiche o di gruppi di persone fisiche» – ovverosia i *place-based systems*.

Restavano invece nella sostanza invariate le lettere *f* e *g*, con la conseguenza che l'unica forma di polizia predittiva ammessa era quella dei *suspect-based systems*. Ciò probabilmente in ragione della loro funzione, come si diceva prima, più spiccatamente "investigativa" – e non preventiva – come tale ritenuta esente dai rilievi di cui sopra.

Erano questi gli opposti poli tra cui ha oscillato l'*AI Act* prima di approdare alla soluzione mediana risultante dal testo definitivo licenziato da Parlamento e Consiglio che sembra ammettere, a determinate condizioni, l'impiego dei sistemi di *predictive policing* e su cui è giunto il momento di concentrare l'attenzione.

Il Regolamento europeo mantiene, all'art. 5, lettera *d*, un divieto rispetto ai sistemi di IA volti a «effettuare valutazioni del rischio relative a persone fisiche al fine di valutare o prevedere la probabilità che una persona fisica commetta un reato, unicamente sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della personalità»[82]. Tuttavia, la stessa disposizione specifica che «tale divieto non si applica ai sistemi di IA utilizzati a sostegno della valutazione umana del coinvolgimento di una persona in un'attività criminosa, che si basa già su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa»[83]. Sono pertanto ammessi i *software person-based* purché impiegati con mera funzione di supporto del processo decisionale della persona fisica, che già si fondi su un ragionevole sospetto, ricadendo in questo caso tra i sistemi ad alto rischio[84].

[82] V. art. 5 reg. 1689/2024/UE.
[83] *Ibidem*.
[84] V. Considerando n. 42: «In linea con la presunzione di innocenza, le persone fisiche nell'Unione dovrebbero

Specularmente, infatti, è stato modificato l'elenco di cui all'Allegato III, paragrafo 6[85] e, in particolare, la nuova lettera *d* ove ora figurano i sistemi utilizzati «per determinare il rischio di reato o recidiva in relazione a una persona fisica non solo sulla base della profilazione [...] o per valutare i tratti e le caratteristiche della personalità o il comportamento criminale pregresso di persone fisiche o gruppi»[86]. Pare dunque che la norma si presti a ricomprendere i sistemi *person-based* che rispettino i "paletti" poc'anzi indicati.

Due sono, a nostro avviso, i principali criteri-guida che hanno condotto alla costruzione dell'attuale assetto normativo e che consentono di tracciare uno spartiacque tra le pratiche vietate e i sistemi ad alto rischio.

Il primo, il rispetto della presunzione di innocenza[87]. La preoccupazione dell'Unione – espressamente manifestata nel Considerando 42 – è quella di assicurare agli individui di essere giudicati per il loro comportamento effettivo. Da qui, l'esclusione del divieto di cui all'art. 5 rispetto ai *person-based systems* che si limitino ad assistere le valutazioni di pericolosità già basate su «fatti oggettivi e verificabili»[88]. Come avremo modo di vedere più nel dettaglio[89], l'ambito di impiego di strumenti di questo tipo potrebbe essere quello delle misure di prevenzione questorili, che tradizionalmente richiedono un giudizio prognostico – *id est* l'accertamento della pericolosità per la pubblica sicurezza – fondato su sospetti.

Il secondo criterio guida è rappresentato dalla funzione di supporto dell'AI alla valutazione umana. I sistemi di AI volti a effettuare prognosi individuali di pericolosità che comportano una profilazione possono essere utilizzati esclusivamente *a sostegno* di un giudizio dell'uomo e mai costituire l'unica base valutativa circa il possibile

---

sempre essere giudicate in base al loro comportamento effettivo. Le persone fisiche non dovrebbero mai essere giudicate sulla base di un comportamento previsto dall'IA basato unicamente sulla profilazione, sui tratti della personalità o su caratteristiche quali la cittadinanza, il luogo di nascita, il luogo di residenza, il numero di figli, il livello di indebitamento o il tipo di automobile, *senza che vi sia un ragionevole sospetto che la persona sia coinvolta in un'attività criminosa sulla base di fatti oggettivi verificabili e senza una valutazione umana al riguardo* [...]» [il corsivo è nostro].

[85] Segnaliamo che è stato modificato anche il titolo del paragrafo 6 che si riferisce ora alle «Attività di contrasto, *nella misura in cui il pertinente diritto dell'Unione o nazionale ne permette l'uso*» [corsivo aggiunto]. La scelta di lasciare (anche) agli Stati membri la facoltà di vietare l'utilizzo di questi sistemi deriva probabilmente dall'esistenza di numerose voci contrarie a queste pratiche, come del resto testimonia la posizione precedentemente assunta dal Parlamento nel giugno 2023.

[86] V. Allegato III, paragrafo 6, lett. *d*.

[87] La presunzione d'innocenza è espressamente richiamata all'*incipit* del Considerando n. 42 reg. 1689/2024/UE, dedicato proprio all'inserimento di alcuni sistemi predittivi nelle pratiche vietate di cui all'art. 5 (v. nota n. 84).

[88] Così il Considerando n. 42 reg. 1689/2024/UE.

[89] V. *infra* § 6.

coinvolgimento di un individuo in una attività criminale. Ciò non solo rappresenta una garanzia fondamentale per il destinatario della predizione, ma testimonia altresì il rifiuto di un atteggiamento di cieco affidamento all'intelligenza artificiale – quella cioè che il Regolamento indica come "distorsione dall'automazione"[90].

Quanto invece ai *place-based systems*, ci sembra che in nessun caso essi possano ricadere nel divieto di cui all'art. 5, in quanto relativo ai sistemi deputati a effettuare "valutazioni del rischio relative a persone fisiche", di cui, come sappiamo, non vi è traccia nella dinamica delle applicazioni di polizia predittiva che mirano all'identificazione delle "zone calde".

Alcuni dubbi potrebbero affiorare rispetto alla collocazione dei sistemi *place-based* tra quelli ad alto rischio. Invero, ciò che nelle precedenti versioni del Regolamento sembrava alludere agli strumenti in parola era il riferimento – "impersonale" – a sistemi orientati a "prevedere il verificarsi o il ripetersi di un reato", che scompare nel testo definitivo. Tuttavia, guardando all'attuale disposto dell'Allegato III e, in particolare, alla seconda parte lettera *d*, si potrebbe far leva sui sistemi utilizzati per valutare il comportamento criminale pregresso e concludere così per la sottoposizione dei *tools* in esame all'articolato reticolo di prescrizioni dell'*AI Act*.

Una simile conclusione – volta, da un lato, a escludere *tout court* il divieto per i *place-based systems* e, dall'altro lato, a considerarli "ad alto rischio" – ci sembra, del resto, trovare conforto sia alla luce dell'essenza stessa dei *place-based systems* sia dalla complessiva lettura del provvedimento.

La prima conclusione si giustifica in quanto i sistemi che individuano gli *hot spot* sono decisamente meno problematici dei *person-based* tanto in ragione dei dati che processano (tipologia, data e orario, luogo di commissione di un reato), i quali non sembrano qualificabili come "personali" – salvo aderire alla interpretazione estensiva sopra illustrata – quanto, e soprattutto, perché essi non hanno effetti diretti sugli individui, limitandosi a razionalizzare gli interventi delle forze dell'ordine in determinate aree urbane con un mero effetto deterrente[91].

Tuttavia – e veniamo alla qualificazione "ad alto rischio" – la prassi statunitense ha dimostrato come le problematiche relative a opacità, qualità dei dati, *feedback loops*, etc. si siano poste anche in relazione a tali sistemi[92], reclamando pertanto il loro

[90] Tale locuzione è contenuta nell'art. 14 reg. 1689/2024/UE, dedicato per l'appunto al requisito della *sorveglianza umana*, di cui si dirà *infra* § 5.
[91] V. *supra* § 2.
[92] V. *supra* § 3.

assoggettamento a determinate misure al fine – chiaramente enunciato nel Considerando n. 59 – di «evitare impatti negativi, mantenere la fiducia dei cittadini e garantire la responsabilità e mezzi di ricorso efficaci»[93]. Peraltro, nell'analisi della disciplina applicabile a tale categoria, è possibile riscontrare alcuni riferimenti impliciti alle criticità proprie dei *place-based systems*; circostanza, questa, che sembra deporre a favore della loro collocazione tra i *tools* di cui all'Allegato III.

L'elenco dei sistemi ad alto rischio rilevante in materia di polizia predittiva si conclude con la (nuova) lettera *e* che, ricalcando la precedente lettera *f*, conferma ancora una volta l'inclusione nella categoria in esame dei sistemi di AI finalizzati alla profilazione nel corso dell'indagine, dell'accertamento e del perseguimento di reati, ovverosia dei *suspect-based systems*.

Inquadrati i diversi sistemi di *predictive policing* nel mosaico del Regolamento europeo, possiamo ora esaminare la disciplina applicabile.

5. I sistemi ad alto rischio – tra cui, come si è visto, rientrano quelli di polizia predittiva ai sensi dell'Allegato III, paragrafo 6 – sono regolati al Capo III dell'*AI Act*, il quale stabilisce i requisiti che devono rispettare (Sezione 2) nonché gli obblighi gravanti sui soggetti che a vario titolo entrano in contatto con gli stessi (Sezione 3).

Innanzi tutto, i fornitori[94] dovranno sottoporre i sistemi ad alto rischio a una *valutazione di conformità*[95] prima dell'immissione sul mercato dell'Unione. Essa è finalizzata a dimostrare che gli stessi rispettino i requisiti fissati dal Regolamento – e di cui diremo subito – nonché a ottenere la "marcatura CE".

In primo luogo, occorrerà istituire un *sistema di gestione dei rischi*[96], costituito da «un processo iterativo continuo pianificato ed eseguito nel corso dell'intero ciclo di vita di un sistema di IA ad alto rischio»[97]. L'obiettivo è quello di identificare e

[93] Così il Considerando n. 59 ove si spiega la ragione alla base della scelta di considerare "ad alto rischio" alcuni sistemi di AI impiegati dalle autorità di contrasto.
[94] V. art. 3, n. 1) reg. 1689/2024/UE, ai sensi del quale il "fornitore" è: «una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che sviluppa un sistema di IA o un modello di IA per finalità generali o che fa sviluppare un sistema di IA o un modello di IA per finalità generali e immette tale sistema o modello sul mercato o mette in servizio il sistema di IA con il proprio nome o marchio, a titolo oneroso o gratuito».
[95] V. art. 8 reg. 1689/2024/UE. Segnaliamo inoltre che l'art. 43 stabilisce che, laddove si tratti di sistemi di IA ad alto rischio di cui all'Allegato III, punti da 2 a 8 – elenco nel quale sono incluse (al punto 6), come visto (v. *supra* § 4), anche le applicazioni di polizia predittiva – i fornitori dovranno seguire la procedura di valutazione della conformità basata sul controllo interno di cui all'Allegato VI, che non prevede il coinvolgimento di un organismo notificato.
[96] V. art. 9 reg. 1689/2024/UE.
[97] Così il Considerando n. 65 reg. 1689/2024/UE.

governare – costantemente – i rischi ragionevolmente prevedibili dei sistemi in esame per la salute, la sicurezza e i diritti fondamentali, attraverso l'adozione di misure accuratamente selezionate dal fornitore stesso – d'intesa, eventualmente, con esperti e portatori d'interessi esterni – dirette alla loro mitigazione.

Un secondo aspetto riguarda *la qualità e la governance dei dati*[98] utilizzati dal sistema ad alto rischio, al fine di garantire un funzionamento sicuro e privo di effetti discriminatori. Pertanto, si richiede che i *set* di dati di addestramento, convalida e prova, siano non solo completi (per quanto possibile) e privi di errori, ma anche "pertinenti" e "sufficientemente rappresentativi" delle categorie di persone nei confronti dei quali il sistema è destinato a essere utilizzato, affinché siano evitate possibili distorsioni – e si aggiunge – «specie laddove gli output di dati influenzano gli input per operazioni future (feedback loops [...])»[99]. Ci sembra questo un interessante, implicito riferimento ai sistemi di polizia predittiva *place-based* che, come si è detto[100], possono dare luogo a questo "circolo vizioso", richiedendo pertanto una specifica considerazione.

Un altro spunto utile per la materia di nostro interesse riguarda i casi – espressamente richiamati – in cui il sistema di AI si basi su "dati storici", i quali «potrebbero essere influenzati da tali distorsioni intrinseche, che sono destinate ad aumentare gradualmente e quindi a perpetuare e amplificare le discriminazioni esistenti, in particolare nei confronti delle persone che appartengono a determinati gruppi vulnerabili, inclusi gruppi razziali o etnici». Evidentemente qui traspare la consapevolezza del legislatore riguardo ai rischi connessi ai sistemi che si alimentano di dati relativi ai reati passati ovvero ai precedenti contatti di un individuo con le forze dell'ordine, che devono essere oggetto di particolare attenzione per evitare forme di discriminazione e risultati non accurati.

Ad ogni modo, per l'intero ciclo di vita del sistema di AI dovrà essere garantito il rispetto del diritto alla vita privata e familiare nonché quello alla *data protection*, trovando applicazione la pertinente normativa eurounitaria – in particolare, rispetto alla polizia predittiva, la c.d. Direttiva *Law Enforcement*[101].

[98] V. art. 10 reg. 1689/2024/UE.
[99] V. Considerando n. 67 reg. 1689/2024/UE.
[100] V. *supra* § 3.
[101] Per un'analisi di tale direttiva nella prospettiva dei suoi riflessi sulla polizia predittiva, volendo v. E. Pietrocarlo, *op. cit.*, 142 ss. e 147 s.

Sul diverso piano della tracciabilità, vengono poi in rilievo le disposizioni relative alla redazione della *documentazione tecnica* [102] e alla *conservazione delle registrazioni*[103], le quali consentono di monitorare in modo continuo il funzionamento del sistema di AI.

Di particolare rilievo, alla luce delle diffuse critiche di opacità mosse ai sistemi di AI, è il tema della *trasparenza*[104]. Al riguardo, si impone che questa sia assicurata "*by-design*", a partire cioè dalla progettazione del *tool* fino al momento in cui lo stesso viene utilizzato dal *deployer*[105], al quale devono essere fornite informazioni chiare e accessibili su capacità, limiti ed esatto *modus operandi*. Solo in questo modo sarà possibile consentire all'utente di assumere "decisioni informate" [106] e sfruttare correttamente l'intelligenza artificiale come valore aggiunto nello svolgimento di mansioni tradizionali. E proprio a questo riguardo il quadro si completa con il requisito della *sorveglianza umana*[107] dei sistemi di AI, quale garanzia del loro corretto e sicuro utilizzo, anche in caso di anomalie di funzionamento o di impatto sui diritti fondamentali. A tal fine, il fornitore dovrà individuare e integrare nel sistema, ove tecnicamente possibile, apposite misure di vigilanza prima della sua immissione sul mercato o messa in servizio, oppure dovrà selezionare in anticipo misure che saranno poi attuate dal *deployer*, il quale dovrà peraltro essere debitamente informato così da sorvegliare consapevolmente il sistema di AI e, se del caso, intervenire.

Infine, i fornitori devono assicurare l'*accuratezza*, la *robustezza* e la *cybersecurity*[108] dei sistemi ad alto rischio. Ciò che si vuole in sostanza evitare è che, a causa di malfunzionamenti dovuti a errori interni (come il più volte citato *feedback loop* dei sistemi *place-based*) ovvero ad attacchi esterni con intenti malevoli (si pensi a *cyberattack*), il sistema di AI possa essere compromesso e, di conseguenza, anche l'accuratezza del suo risultato. Pertanto, è importante incorporare nei sistemi soluzioni tecniche – come, ad esempio, meccanismi di interruzione del funzionamento in presenza di anomalie (c.d. piani *fail-safe*)[109] – e appropriate misure sul piano della

[102] V. art. 11 reg. 1689/2024/UE.
[103] V. art. 12 reg. 1689/2024/UE.
[104] V. art. 13 reg. 1689/2024/UE.
[105] V. art. 3, n. 4) reg. 1689/2024/UE, secondo cui il "*deployer*" si identifica in «una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale».
[106] V. Considerando n. 72 reg. 1689/2024/UE.
[107] V. art. 14 reg. 1689/2024/UE.
[108] V. art. 15 reg. 1689/2024/UE.
[109] V. Considerando n. 75 reg. 1689/2024/UE.

sicurezza cibernetica, che possono corrispondere a quelle eventualmente già implementate sulla base della legislazione dell'Unione in materia di prossima adozione (il riferimento è al c.d. *Cyber Resilience Act*)[110].

A questa nutrita serie di requisiti che mirano ad assicurare la c.d. *trustworthy AI* seguono, come anticipato, obblighi a carico dei soggetti inseriti nel ciclo di vita del sistema di AI. Tra questi, spiccano anzitutto i fornitori, i quali dovranno, *inter alia*, istituire un *sistema di gestione della qualità* [111] che garantisca la conformità al Regolamento dei propri sistemi ad alto rischio, adottare le opportune misure correttive laddove questi non siano *compliant*[112] e collaborare con le autorità competenti[113].

Sebbene si tratti di prescrizioni contenute in un diverso capo del Regolamento[114], riteniamo opportuno dare atto di due ulteriori obblighi, gravanti sui fornitori dei sistemi ad alto rischio, relativi, da un lato, all'istituzione di un *sistema di monitoraggio successivo*[115] alla loro immissione sul mercato al fine di valutare costantemente la conformità del sistema in questione ai requisiti sin qui descritti e, dall'altro lato, alla *condivisione di informazioni su incidenti gravi*[116] con le autorità di vigilanza del mercato degli Stati membri ove essi si sono verificati.

Quanto alle altre categorie di soggetti – rappresentanti autorizzati dei fornitori[117], importatori[118], distributori[119], *deployer*[120] – sono questi ultimi ad assumere peculiare rilievo nell'ottica del legislatore europeo, dal momento che gran parte dei rischi legati ai sistemi di AI deriva proprio dal modo in cui sono utilizzati[121]. Da qui il loro ruolo di garanti – assieme, naturalmente, ai fornitori – della tutela dei diritti fondamentali, in quanto essi, conoscendo più da vicino il contesto di utilizzo del sistema di AI e le specificità dei destinatari, sono in grado di individuare rischi non previsti nella fase di sviluppo. Costoro sono pertanto tenuti ad adottare misure tecniche e organizzative adeguate a un impiego dei sistemi di AI rispettoso del Regolamento nonché a

[110] V. Considerando nn. 77 e 78 reg. 1689/2024/UE.
[111] V. art. 17 reg. 1689/2024/UE.
[112] V. art. 20 reg. 1689/2024/UE.
[113] V. art. 21 reg. 1689/2024/UE.
[114] Il riferimento è al Capo IX, rubricato «Monitoraggio successivo all'immissione sul mercato, condivisione delle informazioni e vigilanza del mercato».
[115] V. art. 72 reg. 1689/2024/UE.
[116] V. art. 73 reg. 1689/2024/UE.
[117] V. art. 22 reg. 1689/2024/UE.
[118] V. art. 23 reg. 1689/2024/UE.
[119] V. art. 24 reg. 1689/2024/UE.
[120] V. art. 26 reg. 1689/2024/UE.
[121] V. Considerando n. 93 reg. 1689/2024/UE.

monitorare il loro funzionamento, assicurare che la persona deputata alla sorveglianza umana disponga delle competenze necessarie[122] e informare le persone soggette alle decisioni del sistema di AI[123]. A tale ultimo riguardo, è interessante richiamare l'art. 86 «Diritto alla spiegazione dei singoli processi decisionali», il quale attribuisce specularmente a qualsiasi persona interessata che sia oggetto di una decisione adottata sulla base dell'*output* di un sistema di IA ad alto rischio elencato nell'Allegato III e che produca effetti giuridici o comunque abbia un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi diritti fondamentali, «il diritto di ottenere dal deployer spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata»[124].

Il quadro degli obblighi gravanti sul *deployer* si completa infine con quello relativo alla *registrazione* – laddove si tratti, come nel caso della polizia predittiva, di organismi di diritto pubblico – in una apposita banca dati dell'Unione nonché quello di svolgere – prima dell'utilizzo iniziale del sistema di AI ad alto rischio e in caso di successivi aggiornamenti – una *valutazione d'impatto sui diritti fondamentali*[125], funzionale a intercettare e gestire prontamente potenziali pregiudizi per gli individui.

Merita infine un accenno il quadro di *governance* e di *enforcement* costruito dal Regolamento.

Sul primo versante, occorre anzitutto distinguere tra gli organismi competenti a livello dell'Unione e quelli a livello dei singoli Stati membri[126].

Nell'architettura europea, è stata anzitutto prevista l'istituzione, da parte della Commissione[127], dell'Ufficio per l'AI[128] con il compito di «sviluppare competenze e

[122] V. Considerando n. 91 reg. 1689/2024/UE.
[123] Segnaliamo che, con particolare riguardo alla polizia predittiva, troverà applicazione la disciplina delle informazioni da rendere disponibili all'interessato di cui all'art. 13 della c.d. Direttiva *Law Enforcement*, il quale è espressamente richiamato dall'art. 26 allorché si tratti di sistemi ad alto rischio utilizzati ai fini di contrasto.
[124] V. art. 86 reg. 1689/2024/UE.
[125] V. art. 27 reg. 1689/2024/UE. Essa consiste in una descrizione dei processi in cui il sistema sarà utilizzato, nonché del periodo di tempo e della frequenza di impiego, delle categorie di persone fisiche e dei gruppi che possono essere interessati dal suo uso, dei rischi per tali soggetti, delle misure di sorveglianza umana e, infine, di quelle da adottare in casi di concretizzazione dei rischi. Il Regolamento si preoccupa altresì di coordinare tale valutazione di impatto con quella prevista nell'ambito della normativa in tema di *data protection* (art. 35 GDPR e art. 27 Direttiva *Law Enforcement*). V. C. Novelli, *op. cit.*, 112, che suggerisce di standardizzare tale valutazione d'impatto affinché sia uniforme per tutti i *deployer* di sistemi AI ad alto rischio.
[126] V. il Capo VII dedicato alla «Governance», la cui Sezione 1 è riservata alla «Governance a livello dell'Unione», mentre la Sezione 2 alle «Autorità nazionali competenti».
[127] V. Decisione della Commissione, del 24.1.2024, che istituisce l'Ufficio europeo per l'intelligenza artificiale C(2024)390.
[128] V. art. 64 reg. 1689/2024/UE.

capacità dell'Unione nel settore dell'IA e di contribuire all'attuazione del diritto dell'Unione in materia di IA»[129]. A questo devono aggiungersi il Comitato europeo per l'IA[130] e il Forum consultivo[131]: il primo, composto da un rappresentante per Stato membro, è responsabile di una serie di compiti consultivi (quali, ad esempio, l'emanazione di pareri, raccomandazioni, consulenze su questioni relative all'attuazione dell'*AI Act*) nei confronti di Commissione, Stati membri e autorità nazionali competenti su questioni specifiche connesse all'IA; il secondo, che rappresenta una selezione equilibrata di portatori di interessi (es. industria, accademia, società civile, etc.), è deputato a fornire consulenza e competenze tecniche al Comitato e alla Commissione. Infine, la Commissione dovrà istituire il Gruppo di esperti scientifici indipendenti[132] competente a sostenere le attività di esecuzione del Regolamento e a supportare, grazie alle competenze scientifiche o tecniche dei componenti selezionati dalla medesima Commissione, l'Ufficio dell'IA.

A livello nazionale, si prevede invece che ciascuno Stato membro istituisca o designi come autorità nazionali competenti del controllo sull'applicazione e l'attuazione del Regolamento almeno un'autorità di notifica e almeno un'autorità di vigilanza del mercato[133].

Sul diverso piano dell'*enforcement*, l'*AI Act* impone agli Stati membri di adottare tutte le misure necessarie, incluse «sanzioni effettive, proporzionate e dissuasive»[134] in caso di violazioni delle prescrizioni in esso contenute, delle quali dovrà essere data apposita comunicazione alla Commissione. Il Regolamento, a seconda del tipo di violazione, fissa delle soglie che hanno riguardo o alla percentuale del fatturato annuo globale della società responsabile nell'anno finanziario precedente ovvero a un importo predeterminato. Tra queste, segnaliamo in particolare la non conformità al divieto di pratiche di IA di cui all'art. 5, assoggettabile a sanzioni amministrative pecuniarie fino a € 35 000 000, se l'autore del reato è un'impresa, fino al 7 % del fatturato mondiale totale annuo dell'esercizio precedente, se superiore, che troverà applicazione, ad esempio, allorché si utilizzino sistemi di polizia predittiva volti a

[129] V. Considerando n. 148 reg. 1689/2024/UE.
[130] V. art. 65 reg. 1689/2024/UE.
[131] V. art. 67 reg. 1689/2024/UE.
[132] V. art. 68 reg. 1689/2024/UE.
[133] V. art. 70 reg. 1689/2024/UE.
[134] V. art. 99 reg. 1689/2024/UE. V. V. Mongillo, *Responsabilità da reato degli enti e crimini connessi all'intelligenza artificiale: tecniche giuridiche di intervento e principali ostacoli*, in www.archiviopenale.it, 19.6.2024, 19 s. per alcune considerazioni sull'ampio margine di apprezzamento lasciato agli Stati membri in relazione al tipo e al contenuto delle sanzioni da adottare nonché sul ruolo da attribuire allo strumento penale in questo settore.

elaborare una prognosi di pericolosità individuale esclusivamente sulla base della profilazione.

Esaurito l'esame dell'ampia congerie di disposizioni dell'*AI Act* rilevanti per la polizia predittiva, non resta che interrogarsi sulla "bontà" delle scelte compiute dal legislatore europeo e sul futuro di questi strumenti nella prassi applicativa.

6. L'analisi sin qui compiuta ci permette ora di verificare se le soluzioni elaborate siano in grado di ovviare alla criticità prima segnalate.

Un primo aspetto su cui ci sembra utile soffermarsi riguarda la posizione in definitiva assunta dalle istituzioni europee in merito all'inquadramento dei sistemi *person-based* e *place-based*, tenuto conto dei cambiamenti di rotta intercorsi, a differenza di quanto accaduto per i *suspect-based systems* sempre saldamente inseriti tra quelli ad alto rischio.

Innanzi tutto, crediamo sia da condividere la decisione di circoscrivere il divieto di cui all'art. 5 ai soli sistemi *person-based* volti a effettuare valutazioni di pericolosità individuali esclusivamente sulla base della profilazione o della valutazione dei tratti e delle caratteristiche della personalità; lo stesso vale, di conseguenza, per la scelta di collocare nella categoria "ad alto rischio" quelli non fondati soltanto su simile trattamento automatizzato. Invero, laddove gli algoritmi predittivi siano impiegati per supportare una valutazione umana circa la pericolosità di un individuo, la quale sia già corredata da sospetti ragionevoli, sembra venir meno il rischio – inaccettabile – di passare da un diritto penale del fatto a un diritto penale d'autore [135], alla base dell'iniziale ripensamento del legislatore europeo. Nella corretta prospettiva indicata dal Regolamento, i sistemi in questione potrebbero rappresentare, in virtù delle loro straordinarie capacità analitiche, un valore aggiunto per l'affidabilità del giudizio prognostico, che potrebbe beneficiare di una base valutativa più consistente. Nel nostro ordinamento, esistono invero istituti che possono essere accostati ai sistemi *person-based* americani che abbiamo prima esaminato, poggiando su un analogo

[135] Tale preoccupazione è sottesa al più volte citato Considerando n. 42 reg. 1689/2024/UE. Nello stesso senso, v. altresì la Risoluzione del 6.10.2021, cit., Considerando Q, ove si sottolinea che «il moderno diritto penale si basa sull'idea che le autorità reagiscono a un reato dopo che è stato commesso, senza supporre che le persone siano pericolose e debbano essere sorvegliate costantemente per prevenire possibili illeciti». In dottrina, v. G. Ubertis, *op. cit.*, 81, che evidenzia, con riguardo ai sistemi predittivi di AI, come «[l]a standardizzazione delle informazioni rischia di far passare dal "diritto penale del fatto" al "diritto penale d'autore"»; v. altresì V. Manes, *op. cit.*, 13 s.; M. Gialuz, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in www.penalecontemporaneo.it, 28.5. 2019, 21, rispetto però all'uso di algoritmi predittivi da parte del giudice penale.

meccanismo di fondo, e rispetto ai quali potrebbe ipotizzarsi l'innesto di questa tipologia di *predictive policing*.

Il riferimento è alle misure di prevenzione questorili dell'avviso orale e dell'ammonimento in materia di *stalking*[136] che, in linea generale, si sostanziano nella diffida (nell'un caso, a non violare legge; nell'altro, a non realizzare atti persecutori) indirizzata nei confronti di un soggetto destinatario di una prognosi di pericolosità[137]. Ed è proprio la realtà – critica – in cui vive l'accertamento prognostico posto a fondamento di queste misure che suggerisce di avviare una riflessione sul ricorso agli strumenti predittivi in discorso, anche in considerazione del "lasciapassare" dell'ultima versione dell'*AI Act* al riguardo. Invero un simile scrutinio nella prassi è rimesso alla pressoché totale discrezionalità del questore[138], in quanto, nonostante gli sforzi di parte della giurisprudenza, non esistono parametri normativi che possano fungere da reale guida[139]. Allora, si potrebbe prevedere che alla valutazione del questore già basata

[136] V. nota n. 69.

[137] Con riferimento all'avviso orale, si tratta della c.d. pericolosità per la pubblica sicurezza: v. F. Menditto, *Le misure di prevenzione personali e patrimoniali*, Milano 2012, 47; in giurisprudenza, v. T.a.r. Veneto Venezia, Sez. III, 27.3.2019, n. 468 che intende tale pericolosità come sospetto «della presenza di elementi tali da ritenere la configurabilità, nel soggetto destinatario dell'avviso, di una personalità propensa a seguire particolari comportamenti antigiuridici». Per quanto riguarda l'ammonimento in materia di atti persecutori, il questore dovrà svolgere un giudizio di pericolosità individuale volto ad appurare la probabilità che il "candidato" realizzi condotte di *stalking* nei confronti della persona offesa che ha richiesto l'applicazione della misura: v. per tutti A.M. Gasparre, *op. cit.*, 315.

[138] V. F. Consulich, *Le misure di prevenzione personali tra Costituzione e Convenzione*, in www.lalegislazionepenale.eu, 18.3.2019, 18; E. Mariani, *Prevenire è meglio che punire. Le misure di prevenzione personali tra accertamento della pericolosità e bilanciamenti di interessi*, Milano 2021, 119. Ciò è confermato anche dal limitato sindacato giurisdizionale: v. T.a.r. Calabria Catanzaro, Sez. I, 7.3.2023, n. 339, secondo cui «l'autorità amministrativa competente [...] gode di ampia discrezionalità nell'accertamento e nella valutazione dei presupposti richiesti dalla legge, ossia dei sospetti, dovendo il sindacato del giudice amministrativo limitarsi solo ad aspetti di manifesta irragionevolezza od arbitrarietà dell'iter logico seguito dall'amministrazione o della motivazione adottata». V. altresì T.a.r. Piemonte Torino, Sez. I, 2.12.2020, n. 791.

[139] Con particolare riguardo alla pericolosità per la pubblica sicurezza, v. F. Basile, *Manuale delle misure di prevenzione*, cit., 77 s.; M. Pelissero, *I destinatari della prevenzione* praeter delictum*: la pericolosità da prevenire e la pericolosità da punire*, in *RIDPP* 2017, 457; A. Martini, *Il mito della pericolosità. Alla ricerca di un senso compiuto del sistema della prevenzione personale*, in *RIDPP* 2017, 547 ss. V. anche A.M. Maugeri, *L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra* evidence based practices *e tutela dei diritti fondamentali*, in www.archiviopenale.it, 17.5.2021, 8. Sulla generale problematicità degli accertamenti prognostici relativi alla propensione dell'individuo alla commissione di reati, previsti in vari segmenti del sistema penale, e sul possibile ricorso ad algoritmi predittivi, v. M. Caianiello, *Dangerous Liaisons. Potentialities and Risks Deriving from the Interaction between Artificial Intelligence and Preventive Justice*, in *European Journal of Crime, Criminal Law and Criminal Justice* 2021, vol. 29, 14 ss. Rispetto all'ammonimento, sono indicativi i risultati della ricerca di E. Mariani, *Le misure di prevenzione personale nella prassi milanese*, in *DPenCont* 2018, 10, 314 s. che, analizzando i provvedimenti applicativi di misure di prevenzione da parte della Questura di Milano tra il 2012 e il 2016 nell'ambito delle Province di Milano e di Monza e Brianza (nonché dal Tribunale di Milano, Sezione Autonoma Misure di Prevenzione), ha osservato come lo scrutinio di pericolosità, oltre a essersi fondato sull'esclusivo

su fondati sospetti si affianchi una verifica mediante un algoritmo predittivo dei dati relativi ai precedenti penali e alle vicende personali dell'individuo, allo scopo di dare maggiore consistenza alla prognosi di pericolosità richiesta ai fini dell'applicazione dell'avviso orale e dell'ammonimento in materia di *stalking*.

Non sembrano peraltro qui ricorrere i problemi in punto di proporzione del trattamento sanzionatorio registratisi, ad esempio, nel Missouri, sebbene anche nell'ambito degli istituti italiani in discorso siano previste conseguenze nei confronti del soggetto insensibile alla precedente diffida. Invero, nel caso dell'avviso orale, queste sono circoscritte alle (numerose ma) specifiche ipotesi tassativamente indicate agli artt. 71, 72 e 73 del Codice Antimafia[140]; quanto all'ammonimento, l'aggravamento di pena e il passaggio alla procedibilità d'ufficio operano soltanto allorché il destinatario realizzi atti persecutori nei confronti della persona che ha richiesto la misura. A ciò si aggiunga che tanto l'avvisato quanto l'ammonito dispongono di diversi mezzi di contestazione del provvedimento emesso a loro carico, compreso l'accesso a un giudice[141].

Quanto ai sistemi *place-based*, allo stesso modo ci sembra condivisibile la loro eliminazione dalle pratiche vietate – nelle quali erano stati inclusi a seguito degli emendamenti presentati dal Parlamento nel 2023 – e la loro conseguente (re)inclusione nel novero dei sistemi ad alto rischio. Ci troviamo infatti al cospetto di sistemi che, come più volte ribadito, non presentano una soglia di potenziale lesività per i diritti fondamentali tale da giustificare un loro divieto. Pertanto, anche alla luce del "via libera" del Regolamento, crediamo sia da favorire l'impiego di strumenti come *X-Law*[142] per una migliore gestione delle forze dell'ordine, consapevoli dell'esperienza

apprezzamento dei precedenti penali del proposto, sia stato rimesso all'intuito del questore, non riscontrandosi criteri guida validi sul piano legislativo.

[140] V. nota n. 69 nonché diffusamente, volendo, E. Pietrocarlo, *op. cit.*, 34 ss.

[141] Quanto all'avviso orale, il destinatario può chiedere in qualsiasi momento la revoca al questore – che opera secondo il meccanismo del silenzio-assenso – mentre, in caso di rigetto, il relativo provvedimento sarà impugnabile mediante ricorso gerarchico al prefetto. Inoltre, l'avvisato potrà presentare ricorso straordinario al Presidente della Repubblica, per soli motivi di legittimità, nonché ricorso al giudice amministrativo, trattandosi di atto amministrativo. V. sul punto F. Menditto, *op. cit.*, 49. Sul diverso versante dell'ammonimento, la giurisprudenza maggioritaria riconosce al destinatario la possibilità di impugnare il provvedimento davanti agli organi di giustizia amministrativa in quanto, in virtù delle sue possibili conseguenze negative, lo stesso è qualificabile come un provvedimento amministrativo immediatamente lesivo: v., *ex multis*, T.a.r. Calabria Reggio Calabria, Sez. I, 4.11.2010, n. 1171; T.a.r. Liguria, 12.1.2010, n. 31. In dottrina, v. in particolare A.M. Gasparre, *op. cit.*, 314; M. D'Arienzo, *La prevenzione del reato di* stalking*. Limiti all'esercizio del potere di ammonimento orale: il sindacato giurisdizionale ed i poteri istruttori del giudice amministrativo*, in *GI* 2012, 11, 2423. In senso contrario, v. T.a.r. Sicilia Palermo, Sez. I, 31.3.2011, n. 605; T.a.r. Sicilia Palermo, Sez. I, 13.4.2010, n. 4957.

[142] Ciò anche alla luce di quanto al tempo dichiarato dall'ideatore di *X-Law* nel corso di un'intervista televisiva

americana e dunque della necessità al contempo di evitare forme di controllo eccessivo su alcuni ambiti territoriali.

Ci convince altresì l'inclusione, tra i sistemi ad alto rischio, delle applicazioni *suspect-based*. Queste – l'esperienza di *KeyCrime* insegna – tendono a essere impiegate a fronte della commissione di diversi reati con caratteristiche di serialità, al fine di offrire un supporto alle tradizionali attività investigative delle forze dell'ordine – piuttosto che, come si diceva, in una logica preventiva –, con risultati apprezzabili in termini di identificazione del profilo del relativo autore finalizzato all'esecuzione di misure coercitive[143]. Il loro minore tasso di problematicità trova del resto conferma nel fatto che durante l'intero *iter* di formazione del provvedimento i sistemi in questione hanno sempre mantenuto loro collocazione tra quelli ammessi, seppure considerati ad alto rischio.

Quanto alla complessiva efficacia del quadro regolatorio riservato ai sistemi *high risky* – categoria al cui interno sono oggi confluiti, come abbiamo cercato di dimostrare, l'intera gamma di strumenti di *predictive policing* – sembra che il reticolo di requisiti stabiliti sia idoneo a minimizzare i problemi sopra illustrati: a partire da quanto prescritto in materia di qualità e *governance* dei dati, di trasparenza, accuratezza, robustezza e cibersicurezza fino alla sorveglianza umana e alla previsione di un sistema di gestione dei rischi nell'intero ciclo di vita del sistema di AI. Allo stesso modo appare importante la previsione di obblighi in capo ai fornitori – tra tutti la valutazione di conformità al Regolamento – e soprattutto ai *deployer*, veri garanti dei diritti fondamentali alla luce della loro prossimità alle persone interessate dai sistemi di AI, i quali dovranno essere accuratamente formati al fine di disporre delle competenze necessarie a "governare" l'AI e, quindi, di assolvere i compiti loro affidati dall'*AI Act*[144].

---

del 29.12.2018 sul positivo impatto del *software* sulla riduzione del tasso di criminalità, pari a una percentuale del -22% nella città di Napoli e del -39% in quella di Prato.

[143] V. lo studio condotto dalla University of Essex: G. Mastrobuoni, *Crime is Terribly Revealing: Information Technology and Police Productivity*, in *The Review of Economic Studies* 2020, vol. 87, n. 6, 2727 ss. V. inoltre i dati riportati da M. Gialuz, S. Quattrocolo, *op. cit.*, 6.

[144] Questo aspetto è peraltro sottolineato in una delle Risoluzioni – segnatamente la n. 8 – elaborate dall'*Association Internationale de Droit Pénal* all'esito dell'*International Colloquium* tenutosi a Buenos Aires (28-31.3.2023) su *Artificial Intelligence and Administration of Criminal Justice* in vista del *XXI International Congress of Penal Law on Artificial Intelligence and Criminal Law* svoltosi a Parigi (25-28.6.2024): «States and law enforcement authorities must ensure that their personnel who use AI to prevent, detect, or investigate criminal offenses receive hands-on training in the proper use of the relevant AI system, as well as training with respect to the risk of error and bias. They must ensure that such personnel have a thorough knowledge of the dangers AI may pose to human rights». Il testo completo delle Risoluzioni è contenuto nel volume n. 2/2023 della *Revue*

Queste previsioni si rivelano funzionali anche a un impiego degli strumenti in questione nell'ambito delle misure di prevenzione questorili prima menzionate, assicurando il rispetto dei diritti in gioco. Ciò in quanto, sulla base delle regole stabilite nell'*AI Act*, si imporrebbe anzitutto al fornitore di effettuare la suddetta valutazione di conformità, che attesti, ad esempio, la rappresentatività del *set* di dati utilizzato, la trasparenza circa il *modus operandi* del *software*, la predisposizione di misure di sorveglianza umana, etc. Il che consentirebbe di ovviare alle problematiche riscontrate nella prassi e relative principalmente a debolezze "intrinseche" agli algoritmi (qualità dei dati, trasparenza, accuratezza, etc.).

Inoltre, la stessa autorità di pubblica sicurezza, in quanto *deployer*, dovrà anzitutto eseguire una valutazione d'impatto sui diritti fondamentali nonché adottare le misure tecniche e organizzative necessarie ad assicurare il rispetto del Regolamento e a monitorare il funzionamento del sistema, così come sarà tenuta a fornire ai destinatari della predizione algoritmica tutti i chiarimenti necessari sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi posti a fondamento dell'esito adottato. In definitiva, tale articolato *corpus* di prescrizioni appare idoneo a garantire il proficuo e legittimo utilizzo delle tecnologie predittive *person-based* in discorso.

Anche per le altre categorie di sistemi l'*AI Act* sembra rappresentare un'opportunità nel senso appena indicato. Peraltro, con particolare riguardo alle criticità mostrate dai *place-based systems* nelle loro principali modalità operative (il c.d. *feedback loop*), il Regolamento prescrive la predisposizione di «adeguate misure di attenuazione»[145].

In definitiva, ci sembra di poter esprimere un giudizio positivo sul modo in cui l'*AI Act* affronta il tema della polizia predittiva. Ad ogni modo, si tratta di un giudizio necessariamente provvisorio, dovendo essere rinviato alla successiva ed effettiva implementazione del Regolamento, le cui tempistiche si prospettano alquanto dilatate[146]. Peraltro, non è del tutto remoto il rischio che la velocità dell'evoluzione tecnologica renda la legislazione non più attuale, dal momento che le regole dei sistemi

Internationale de Droit Pénal, p. 383 ss.

[145] V. art. 15 reg. 1689/2024/UE.

[146] L'implementazione del Regolamento avverrà in maniera graduale, prevedendosi che: sei mesi dopo l'entrata in vigore, gli Stati membri saranno tenuti a eliminare i sistemi di AI vietati di cui all'art. 5; dopo dodici mesi, troveranno applicazione gli obblighi relativi alla *governance* dell'AI per finalità generali; dopo ventiquattro mesi il Regolamento sarà pienamente applicabile, compresa la parte relativa agli obblighi per i sistemi ad alto rischio di cui all'Allegato III (in cui ricade la polizia predittiva); l'unica eccezione è rappresentata dagli obblighi per i sistemi ad alto rischio definiti nell'Allegato II (elenco della normativa di armonizzazione dell'Unione) che saranno operativi dopo trentasei mesi dall'entrata in vigore dell'atto. V. Camera dei Deputati, *Il regolamento UE in materia di intelligenza artificiale*, 5.2.2024, 13.

ad alto rischio si applicheranno trascorsi due anni dall'entrata in vigore del provvedimento [147]. Inoltre, occorrerà verificare anche la posizione che sarà concretamente assunta da ciascuno Stato membro rispetto alla legislazione in materia di AI nonché, in particolare, di *predictive policing*, dal momento che i sistemi di AI destinati a essere impiegati nelle attività di contrasto sono ammessi nella misura in cui – come esplicita la rubrica del paragrafo 6 dell'Allegato III del Regolamento che li elenca – «il pertinente diritto dell'Unione o nazionale ne permette l'uso».

Tra gli altri, un aspetto di sicuro rilievo che dovrà essere affrontato al fine di assicurare l'effettività del quadro regolatorio in esame riguarderà la scelta circa il tipo di sanzioni da stabilire ad esempio nel caso – rilevante ai fini della nostra analisi – di violazione del divieto di pratiche vietate di cui all'art. 5[148].

Volgendo lo sguardo all'Italia, è in cantiere un disegno di legge[149] che – oltre a intervenire sul diverso piano della criminalizzazione delle condotte realizzate attraverso sistemi di AI[150] – fissa una serie di principi generali sull'uso dell'AI che ricalcano in gran parte quelli sanciti a livello UE[151] e delega altresì il Governo ad adeguare la normativa nazionale all'*AI Act* nonché a definire organicamente la disciplina nei casi di utilizzo di sistemi di intelligenza artificiale per finalità illecite[152]. Inoltre, la bozza contiene disposizioni di dettaglio in relazione ai diversi settori in cui l'AI può essere impiegata tra cui quello – contiguo a quello di nostro interesse –

[147] V., sul diverso tema dell'urgenza di regolare il *novum*, le considerazioni G. Finocchiaro, *Intelligenza artificiale. Quali regole?*, Bologna 2024, 4 ss.

[148] V. F. Mirò-Llinares, *General Report*, in *Revue Internationale de Droit Pénal* 2024, 1, 47 – redatto all'esito dell'*International Colloquium* dell'*Association Internationale de Droit Pénal* tenutosi a Bucharest (14-16.6.2023) su *Criminalisation of AI-related offences* in vista del *XXI International Congress of Penal Law on Artificial Intelligence and Criminal Law* svoltosi a Parigi (25-28.6.2024) – il quale riconosce la possibilità di introdurre una apposita fattispecie di reato, specificando tuttavia che, alla luce del principio di sussidiarietà o *extrema ratio* dell'intervento penale, la stessa non dovrebbe poggiare esclusivamente sulla violazione del divieto di cui all'art. 5 reg. 1689/2024/UE, dovendosi piuttosto richiedere «either a harmful result or a probable risk of harm».

[149] V. d.d.l. n. 1146 recante «Disposizioni e delega al Governo in materia di intelligenza artificiale», attualmente sottoposto all'esame del Senato.

[150] L'art. 25, co. 1, lett. *e* del d.d.l. mira a introdurre una fattispecie di reato – il nuovo art. 612-*quater* Cp rubricato «Illecita diffusione di contenuti generati o manipolati con sistemi di intelligenza artificiale» – volta a reprimere il fenomeno dei c.d. *deep fake*, nonché, ai sensi dell'art. 25, co. 1, lett. da *a* a *d*, una circostanza aggravante comune e altre speciali che poggiano sulla commissione di reati mediante l'impiego di sistemi di AI. Per una panoramica sulle possibili interferenze dei sistemi di AI nella realizzazione di reati, sull'adeguatezza delle fattispecie vigenti a "coprire" tali nuove forme di criminalità e sulle prospettive *de jure condendo*, v. A. Gullo, R. Flor, *Italian Report on Criminalisation of AI-related Offences*, in *Revue Internationale de Droit Pénal*, 2024 1, 190 ss.

[151] V. artt. 3, 4 e 5 del d.d.l.

[152] Le deleghe sono contenute nell'art. 22 del d.d.l.

dell'attività giudiziaria[153], rispetto al quale si stabilisce che «[i] sistemi di intelligenza artificiale sono utilizzati esclusivamente per l'organizzazione e la semplificazione del lavoro giudiziario nonché per la ricerca giurisprudenziale e dottrinale», con l'ulteriore specificazione che «[è] sempre riservata al magistrato la decisione sulla interpretazione della legge, sulla valutazione dei fatti e delle prove e sulla adozione di ogni provvedimento». Sembrerebbe pertanto – *prima facie* e (almeno) per il momento – esclusa la possibilità di ricorrere ad algoritmi predittivi, ad esempio, nella fase di commisurazione della pena o in quella cautelare per la valutazione delle omonime esigenze come al contrario avviene nell'ordinamento nordamericano[154].

Quanto alla polizia predittiva, non si scorge di contro alcun riferimento, sicché, salvo diverse modifiche nel corso dell'*iter* parlamentare ovvero nuovi provvedimenti in materia, deve allo stato escludersi il profilarsi di un divieto rispetti al ricorso a questi sistemi. Sarebbe ad ogni modo auspicabile una precisa presa di posizione da parte del legislatore volta a regolamentare l'uso di siffatti strumenti, prima di tutto attraverso l'istituzione di un meccanismo autorizzativo preventivo[155].

Per le ragioni che confidiamo di aver illustrato, crediamo in definitiva che l'impiego degli strumenti in esame – sempre che ovviamente avvenga nel pieno rispetto dei limiti ad oggi fissati nell'*AI Act* e di quelli che saranno in futuro introdotti nel solco della legislazione eurounitaria – possa rappresentare una opportunità non solo per rafforzare le capacità investigative e preventive delle forze dell'ordine ma anche per dotare di maggiore consistenza i giudizi prognostici affidati all'autorità di pubblica sicurezza.

[153] V. art. 24 del d.d.l.

[154] Sul tema, v. per tutti S. Quattrocolo, *Artificial Intelligence, Computational Modelling and Criminal Proceedings. A Framework for A European Legal Discussion*, Cham 2020, 132 ss.

[155] In questo senso si esprime la prima delle Risoluzioni elaborate dall'*Association Internationale de Droit Pénal* prima citate: «[u]se of AI systems by public authorities for assistance when preventing, detecting, or investigating criminal offences must be authorized in advance by a law or a norm of equivalent binding force». V. il volume n. 2/2023 della *Revue Internationale de Droit Pénal*, cit., 383.