

Polícia, tecnologia i la “revolució” de la intel·ligència artificial¹

JOSÉ MARÍA LÓPEZ RIBA
josemaria.lopez@udg.edu

Universitat de Girona

Resum

L'ús de la intel·ligència artificial per part de la policia ha generat grans expectatives, presentant-se com una revolució tecnològica. No obstant això, la seva incorporació al treball policial no implica un canvi profund en el paradigma, sinó més aviat una automatització de tasques específiques. Així, la IA pot ser útil per a certes funcions, però no transforma fonamentalment la naturalesa del treball policial, com ho van fer altres innovacions tecnològiques anteriors. A més, les promeses de major precisió, i amb això major eficiència, dels models predictius basats en IA són sovint exagerades. La complexitat dels models d'aprenentatge automàtic no garanteix necessàriament millores en la precisió. Així mateix, molts dels sistemes utilitzats, per a la predicció de la criminalitat per exemple, no tenen fonaments científics sòlids i poden perpetuar prejudicis, posant en dubte també les promeses sobre la reducció del biaix humà. El debat públic sobre l'ús de la IA a la policia es complica per la manca de transparència i rendició de comptes, i per la tecnificació del debat. En lloc de centrar-se només en aspectes tècnics, és essencial discutir

les implicacions ètiques, socials i democràtiques de l'ús d'aquestes eines. L'adopció de la IA en la policia pot acabar sent una solució aparent si no hi ha abans una reflexió profunda sobre el seu impacte real.

Paraules clau: policia, tecnologia, automatització, intel·ligència artificial.

1. Una versió anterior d'aquesta recerca es va publicar a López Riba, J. M. (2024). Inteligencia artificial i control policial: cuestiones para un debate frente al hype. *InDret: Revista para el análisis del Derecho*, (2), 407-436. <https://doi.org/10.31009/InDret.2024.i2.10> Vull agrair a la Directora, la Dra. Elena Larrauri, i el subdirector, el Dr. Daniel Varona, de la secció de Criminologia de la revista, la seva amabilitat en permetre reutilitzar algunes parts de la citada publicació.

1. La incorporació d'innovacions tecnològiques al treball policial

La policia, per dur a terme les seves funcions, sempre s'ha recolzat en algunes innovacions tecnològiques. En l'actualitat, quan pensem en tecnologia, pensem principalment en intel·ligència artificial (IA) o robòtica. Però cal destacar que les tecnologies que més impacte han tingut en la forma en què la policia fa la seva feina no han sigut ni serien segurament objecte de novel·les de ciència-ficció. Per exemple, cap altra tecnologia com l'automòbil, el telèfon o la ràdio han suposat una transformació tan significativa del dia a dia policial. Posteriorment, la informàtica, les càmeres de videovigilància i els sistemes d'informació geogràfica també van suposar alguns canvis en la feina a les comissaries de policia.

La introducció del telèfon, la ràdio bidireccional i l'automòbil, a principis del segle XX, van suposar canvis profunds en la forma de treballar dels agents de policia i en la comunicació amb la resta de la societat. Abans del telèfon, els ciutadans havien de recórrer a la policia en persona, desplaçant-se físicament per trobar-los a dependències policials o al carrer. La introducció i generalització de l'ús de la telefonia va obrir un canal de contacte directe i va fer accessible la policia en qualsevol lloc i moment, la qual cosa va provocar també un augment significatiu de les demandes de servei. Quant a la ràdio bidireccional, va proporcionar un mitjà de comunicació instantani entre els comandaments i els agents, així com un canal constant d'intercanvi d'informació entre aquests últims. Anteriorment, els agents havien de dependre de casetes telefòniques o tornar a la comissaria per comunicar-se, cosa que limitava la seva capacitat de resposta. La ràdio va permetre una coordinació més efectiva de les respostes policials, cosa que va permetre el fet que els agents fossin enviats directament als llocs dels incidents. Per la seva banda, l'automòbil va transformar la mobilitat policial, i això va permetre als agents el fet de patricular àrees més extenses amb més rapidesa i respondre més ràpidament a les trucades. Amb cotxes, la policia va poder cobrir zones que abans eren inaccessibles a peu o amb cavall, la qual cosa també implicava un augment de la visibilitat policial (Stroshine, 2015).

En conjunt, aquestes tres innovacions tecnològiques van marcar un punt d'inflexió i van contribuir significativament a la formació de la policia moderna tal com la coneixem avui dia. De fet, es podria afirmar que aquestes tecnologies van donar peu al fet que avui en dia es pugui dir que una de les característiques que defineixen la policia sigui la seva "exhaustiva ubicació espai-temporal (...) i la relació constant que manté amb tota la població" (Sabaté Delgado, 1984, p. 134).

Una altra innovació tecnològica que va ser molt útil per a les organitzacions policials va ser la informàtica, concretament la capacitat de generar registres digitals. La policia sempre s'ha caracteritzat per ser una organització que recull i genera molta informació (Manning, 2008). Tradicionalment, aquesta informació es registrava de forma manual o amb màquina d'escriure, però la incorporació d'equips informàtics va permetre la seva digitalització. Això va comportar una major eficiència en la gestió de la informació (cerques d'informació més ràpides) i una major capacitat per compartir-la (entre unitats o entre diferents organitzacions) (Stroshine, 2015).

Tot i que no són usades principalment per la policia, sinó més aviat per empreses o administracions, la implementació de càmeres de videovigilància van suposar noves oportunitats per a la policia. La seva presència en espais públics i privats ha permès a les organitzacions policials obtenir proves visuals per investigar delictes, identificar sospitosos i comprovar testimonis de manera més ràpida (Goold, 2004).

La idea de "mapejar" els incidents no era nova a les organitzacions policials quan es van introduir els Sistemes d'Informació Geogràfica, però fins llavors s'havia fet de forma rudimentària: amb mapes impresos i xinxetes. Amb la nova tecnologia es podia georeferenciar més informació (no només delictes, sinó també altres incidents o les mateixes actuacions policials) i de forma més precisa. Actualment, l'anomenat *crime mapping* ("mapes de delinqüència") és una de les tècniques més utilitzades a l'hora de distribuir els recursos policials (Manning, 2008; Stroshine, 2015).

2. La “revolució” de la intel·ligència artificial i les seves promeses en el camp de la seguretat

Per parlar de les aplicacions de la IA i les seves promeses en l'àmbit policial primer cal parar atenció al context en el qual es dona aquesta “innovació”. Posteriorment, es parlarà específicament de les eines concretes que fan ús d'aquesta tecnologia. I, finalment, es profunditzarà en els estudis que s'han dut a terme sobre la seva aplicació.

2.1. EL CONTEXT

Cal tenir en compte el context en el qual s'han introduït aquestes eines. Concretament, es podria posar el focus en quatre elements: les expectatives (exagerades), els interessos econòmics darrera d'aquestes, el marc mental del tecnosolucionisme i el context concret en l'àmbit policial.

La idea d'una Quarta Revolució Industrial impulsada per tecnologies com la IA (també la robòtica avançada i el *Big Data*) ha generat debat. Mentre que algunes personalitats influents, com Klaus Schwab² (president executiu del Fòrum Econòmic Mundial) o James Manyika³ (vicepresident de Google) la consideren una realitat, hi ha autors que sostenen que aquestes tecnologies són principalment evolucions de les ja presents a la Tercera Revolució Industrial (processadors, Internet, telèfons mòbils, etc.) i que no hi ha prou evidència de transformacions socials i econòmiques per a considerar-la una nova revolució (p. e., Moll, 2021; 2023). Tot i la gran quantitat d'expectatives i la propaganda al voltant de la IA, moltes de les seves aplicacions actuals no funcionen com es promet i, de fet, algunes no poden funcionar en absolut. Però aquestes expectatives exagerades persisteixen i fan sobreestimar les capacitats de la IA (Narayanan i Kapoor, 2024).

Les empreses que desenvolupen i venen productes d'IA tenen un gran incentiu per exagerar les

capacitats de la seva tecnologia per atraure inversors i clients. A més, els investigadors, sovint pressionats per obtenir finançament i publicar resultats impactants, contribueixen a la creació d'aquesta bombolla d'expectatives. Aquesta situació es veu agreujada per la superficialitat de la cobertura mediàtica, que sovint prioritza els titulars cridaners i les històries sensacionalistes sobre anàlisis rigoroses i crítiques. La manca d'explicacions detallades sobre les limitacions de la IA i la tendència a reproduir sense qüestionar les afirmacions grandiloqüents de les empreses i els investigadors contribueixen a la perpetuació dels mites i creen una falsa imatge de les capacitats reals de la IA (Narayanan i Kapoor, 2024).

S'ha de destacar que aquesta publicitat que sobreestima les capacitats reals de la IA encaixa molt bé en el marc mental del *tecnosolucionisme*, la creença que la tecnologia pot solucionar tots els problemes de la societat, inclosos els més complexos. Una idea perillosa que amaga els interessos econòmics de les empreses encarregades de vendre les tecnologies i que simplifica els problemes socials, desplaçant solucions que sí que tenen en compte les complexitats de la vida social (Mozorov, 2015).

És en aquest context general en què s'ha d'entendre la introducció de la IA en el treball policial. Però abans de veure les aplicacions concretes actuals, cal parar atenció a una eina que es podria considerar com a precursora del *predictive policing* (“polícia predictiva”). Es parla de CompStat (abreviació de Compare Stats). Aquesta eina és un sistema de gestió policial que es va implementar a Nova York als anys 90 per millorar l'eficiència en el control policial. Es basava en la recopilació i l'anàlisi de dades sobre delictes per identificar patrons i tendències en diferents zones de la ciutat. Aquest sistema permetia als comandaments policials veure on es concentra-

2. Veure: <https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab/>

3. Veure: https://elpais.com/tecnologia/2023-06-28/james-manyka-google-la-inteligencia-artificial-cambiara-el-mundo-como-las-computadoras-o-la-electricidad.html?utm_medium=social&utm_campaign=echobox&utm_source=Twitter&ssm=TW_CM_TEC#Echobox=1687935623



ven els delictes, amb quina freqüència es donaven i quin tipus d'incidentes eren més comuns en cada àrea. Però CompStat era principalment una eina de gestió interna i rendició de comptes: la informació generada es feia servir en reunions setmanals, on els caps de districte havien de defensar els seus "resultats" davant els comandaments (Benbouzid, 2019; Moore, 2003).

Si bé CompStat ha estat considerat com una innovació administrativa important en la policia, el seu impacte en l'àmbit de la innovació tecnològica és més subtil, donat que en aquest sentit va representar abans una innovació en la implementació i l'aplicació de la tecnologia, més que en el desenvolupament tecnològic en si. En lloc d'introduir noves tecnologies, va aprofitar la tecnologia existent, particularment els registres digitals i els Sistemes d'Informació Geogràfica, per reformar els sistemes de gestió i responsabilitat policial. Amb aquest sistema, la transició digital va permetre als departaments de policia recopilar, analitzar i disseminar informació sobre incidents i actuacions policials de manera més ràpida, precisa i detallada (Moore, 2003).

CompStat i les tècniques policials relacionades representen bé el gir preventiu del model policial (recolzat sobre la tecnologia), l'anticipació als comportaments infractors per sobre del seu paper tradicionalment reactiu (San Martín Segura, 2020), i la introducció del actuarialisme en la feina de la policia, és a dir, la introducció de saber tècnics dirigits a la predicció de riscos (Harcourt, 2007).

William Bratton, com a cap de la Policia de Nova York, va ser la figura encarregada de la implementació de CompStat. A final de la dècada dels 2000, en un context de crisi econòmica i retallades en el pressupost de les organitzacions policials, i ara com a cap de la Policia de Los Angeles, Bratton va començar a treballar amb agències federals i acadèmics per tal de desenvolupar noves eines més eficients en la gestió dels recursos policials, que van donar lloc als primers models de policia predictiva (Benbouzid, 2019; Brayne, Rosenblat i Boyd, 2015; Ferguson, 2017).

En resum, les aplicacions de la IA en la feina policial es donen en un context general d'expectatives

exagerades sobre les seves capacitats i un clima favorable en l'àmbit policial a la innovació tecnològica, la feina preventiva i la predicció de riscos.

2.2. LES APLICACIONS DE LA INTEL·LIGÈNCIA ARTIFICIAL EN EL TREBALL POLICIAL

Les eines basades en IA aplicades al control policial es poden classificar segons la tecnologia que les sustenta. Una proposta de classificació seria separar-les en tres categories: les eines de reconeixement de patrons, les de predicció de riscos o delictes i les de verificació específica o anàlisi focalitzada.

En primer lloc, les eines de reconeixement són aquelles que analitzen dades existents per identificar coincidències, patrons o trets específics que permetin monitorar situacions o individus en temps real. Per exemple, el reconeixement facial i el reconeixement de matrícules entren dins d'aquesta categoria, ja que utilitzen dades biomètriques o visuals per identificar persones o vehicles. El reconeixement facial implica el processament biomètric en temps real d'imatges captades per càmeres de vídeo fixes o mòbils, amb l'objectiu de trobar coincidències amb bases de dades policials. Quan l'algoritme detecta una possible correspondència, envia una alerta que és revisada per un agent de policia (Fussey, Davies i Innes, 2021). A Espanya, ja s'han implantat càmeres de videovigilància amb reconeixement facial en espais públics com estacions d'autobusos (Miralles, Campisi i Díaz, 2021) o recintes d'esdeveniments com l'Ifema de Madrid (Calatayud, 2021). Tot i això, aquests sistemes no són populars a Europa i, de fet, la regulació de la IA aprovada prohibeix el seu ús *a priori* (deixant obert la seva utilització en casos molt concrets)⁴.

El reconeixement de matrícules consisteix en càmeres instal·lades a les carreteres que processen en temps real les matrícules dels vehicles i les comparen amb bases de dades policials o administratives. Són els sistemes de vigilància basats en IA més estesos a Europa (Jansen, 2018). Estan implementats, per exemple, a la zona de baixes emissions de Barcelona (Miralles et al., 2021). En aquesta categoria també s'hi poden

4. Veure: <http://data.europa.eu/eli/reg/2024/1689/oj>

incloure sistemes que analitzen grans volums de textos en xarxes socials, com per exemple el del Ministeri de l'Interior per detectar discursos d'odi⁵. Aquests tipus de sistemes utilitzen algoritmes que entrenen models capaços de processar, analitzar i interpretar textos de manera automàtica per identificar contingut que pugui incitar a l'odi, la violència o la discriminació en línia mitjançant tècniques de processament de llenguatge natural. Requereixen un “entrenament” amb exemples classificats com llenguatge d'odi (MacAvaney et al., 2019). Tot i que la vigilància policial fa temps que es recolza en la tecnologia, com s'ha vist anteriorment, la incorporació de la IA en aquest àmbit hauria de possibilitar el reconeixement, la recollida i el processament de dades en temps real (Jansen, Sánchez-Monedero i Dencik, 2021).

En segon lloc, les eines predictives es basen en l'anàlisi de dades per anticipar esdeveniments futurs. Sovint es coneixen com a *predictive policing* (“polícia predictiva”) i es poden dividir en dues subcategories: les eines que prediuen on i quan és probable que es cometi un delictes (*place-based*) i les que calculen el risc associat a individus concrets (*person-based*). Aquestes últimes són menys utilitzades per la seva falta de fonaments teòrics i les seves implicacions per a la presumpció d'innocència (Miró-Llinares, 2020; Shapiro, 2017), ja que els falsos positius són molt més greus en aquest tipus d'eines que en les que calculen el risc vinculat a determinats llocs (Ferguson, 2017). Els primers i més utilitzats, els sistemes *place-based*, funcionen en tres etapes. En primer lloc, es recullen dades com ara el tipus de delictes, l'hora exacta de l'incident i la localització específica. En segon lloc, aquestes dades són processades per algoritmes amb l'objectiu de predir la probabilitat que es produeixin delictes en una àrea d'interès. Finalment, la policia utilitza aquestes prediccions per dissenyar i planificar les seves estratègies d'actuació (Brantingham, 2017). Aquestes prediccions permetrien a la policia planificar millor la seva presència al territori i orientar les seves estratègies. Tot i que aquestes eines s'han utilitzat especialment als Estats Units, s'està estudiant la seva implantació en al-

tres països, amb l'expectativa que el seu ús creixi considerablement en el futur⁶ (Bennet Moses i Chan, 2018; Miró Llinares, 2020; Shapiro, 2017). A Espanya, podem trobar exemples com el sistema Pred-Crime de la Policia Local de Rivas-Vaciamadrid (Calatayud, 2021).

Finalment, hi ha les eines de verificació o anàlisi focalitzada, que estan pensades per avaluar casos concrets o validar dades específiques. Aquestes eines no pretenen fer vigilància massiva ni anticipar esdeveniments, sinó que es dediquen a analitzar informació ja recopilada. Un exemple d'implementació d'aquests tipus d'eines al territori espanyol és VeriPol, un sistema que utilitza l'anàlisi de llenguatge natural per determinar si una denúncia de robatori pot ser falsa (Miralles et al., 2021). Aquest software es basa en patrons lingüístics identificats en denúncies anteriors, és a dir, patrons de denúncies identificats com a falsos per la mateixa policia (Martinez Garay et al., 2022).

Tenint en compte el context exposat anteriorment, aquestes eines aplicades al treball policial van acompanyades de les promeses de major precisió en la detecció/predicció, mitjançant tècniques d'aprenentatge automàtic i de processament del llenguatge natural, la major eficiència en la distribució de recursos (com a conseqüència, en part, d'una major precisió) i la major “justícia” en les decisions, en reduir el biaix humà, en la lluita contra la delinqüència (Bennet Mosses i Chan, 2018; Brayne et al., 2015; Fussey et al., 2021; Martinez Garay et al., 2022; Shapiro, 2017; 2019; Wang, Kapoor, Barocas i Narayanan, 2024).

2.3. QUÈ DIU LA RECERCA SOBRE LES PROMESSES QUE ACOMPANYEN AQUESTES EINES?

Tal com sosté Varona Gómez (2024), aquestes afirmacions haurien de ser validades mitjançant l'anàlisi empírica i no acceptades de forma acrítica. Ja existeix recerca, tot i que potser no suficient al nostre context, per donar resposta

5. Veure: https://elpais.com/tecnologia/2018/11/01/actualidad/1541030256_106965.html

6. Tot i que a Europa se esta limitant, per exemple a partir de la regulació esmentada (veure nota 4), i a altres contextos també hi han corrents per aturar o revertir la seva implementació. Veure, per exemple: <https://www.wsj.com/articles/after-backlash-predictive-policing-adapts-to-a-changed-world-11625752931>

a aquestes promeses. Per tal de respondre, es farà una revisió d'algunes de les recerques que permeten avaluar si amb la IA s'augmenta la precisió, l'eficiència i la justícia en la funció del control del delictes de la policia.

2.3.1. MAJOR PRECISIÓ

Pel que fa als sistemes de reconeixement i la seva major precisió, no hi ha molta recerca específica sobre la precisió dels utilitzats per les organitzacions policials. El problema és que no està molt clar quins criteris s'han d'utilitzar per avaluar la seva efectivitat: si el nombre d'identificacions precises, o minimitzar els falsos positius (les identificacions imprecises), etc. (Fussey et al., 2021). En els resultats de la investigació de Fussey et al. (2021) es van identificar diferències en la precisió dels sistemes utilitzats per la policia del sud de Gal·les i la Policia Metropolitana de Londres, arribant a la conclusió que el factor determinant no era el sistema en si mateix, sinó la manera com es desplega, ja que la tecnologia era la mateixa. Hi ha estudis que assenyalen que la precisió dels sistemes de reconeixement facial està molt influenciada i, per tant, varia molt segons les característiques físiques i demogràfiques de les persones (Purshouse i Campbell, 2022), punt que s'exemplificarà en apartats posteriors. Amb els sistemes de reconeixement de trets biomètrics, en general, existeix un problema amb els falsos positius. Adensamer i Klausner (2021) anomenen aquest problema com la "fal·làcia de la taxa base". Cap sistema d'aparellament pot garantir una taxa d'èxit del 100% i, a mesura que augmenta la mida dels conjunts de dades, també creix el risc d'errors, especialment els falsos positius. Això passa perquè, en treballar amb bases de dades més grans, hi ha més probabilitats que es generin coincidències incorrectes. Si, a més, els models estan configurats o "forçats" per trobar coincidències, és probable que acabin identificant resultats que no són precisos, sacrificant la qualitat de les prediccions per complir amb aquest objectiu.

Respecte als sistemes predictius, concretament als *place-based*, hi ha avaluacions que assenyalen una major precisió que els agents de policia o analistes de dades. Per exemple, el sistema

predictiu de l'empresa Geolítica (abans coneguda com a Predpol), segurament el més famós d'aquests sistemes, fa ús d'un algoritme (ETAS, *Epidemic Type Aftershock Sequence*) que entrena un model amb més precisió que els analistes amb les tècniques tradicionals (Mohler et al., 2015). Però posteriors desenvolupaments posen en dubte la suposada major precisió d'aquest sistema. El sismòleg David Marsan, desenvolupador original de l'algoritme ETAS, va ser convidat per posar a prova la precisió d'aquest sistema en comparació amb estratègies més tradicionals, com ara els mapes de distribució de delictes⁷, utilitzant dades obertes de la ciutat de Chicago. Marsan va qüestionar que l'ús de l'algoritme aportés cap avantatge significatiu respecte als mètodes clàssics. Segons la seva anàlisi, la contribució de l'efecte contagi —principi fonamental darrere de PredPol— en la predicció de futurs delictes és molt limitada. A més, va assenyalar que l'assumpció que els delictes futurs seguiran una dinàmica similar als del passat (l'anomenada no estacionalitat) resulta problemàtica, especialment perquè, a diferència dels fenòmens naturals com els terratrèmols, el comportament humà pot canviar, per exemple en resposta a noves estratègies policials (Benbouzid, 2018). A més, una investigació periodística recent va revelar que les prediccions generades pel programari Geolítica a la ciutat de Plainfield tenien una taxa d'encert aproximada de només l'1%⁸.

Altres estudis han valorat l'eficàcia, és a dir, si les estratègies policials basades en les prediccions d'aquests tipus de sistemes han anat acompanyades de reduccions en delictes. Ratcliffe et al. (2021) van comparar els resultats del patrullatge tradicional amb el patrullatge guiat per programari predictiu. Aquests mostren que, en aquelles àrees on s'utilitzava el sistema predictiu, es van observar reduccions significatives en els delictes contra la propietat. D'altra banda, també hi ha avaluacions de l'eficàcia d'aquests sistemes que no obtenen resultats tan positius. Per exemple, Hunt, Saunders i Hollywood (2014) van avaluar la implementació d'un sistema predictiu per part de la policia de Shreveport. En la seva avaluació, van trobar que, tot i que els agents percebien beneficis en l'ús d'aquestes noves estratègies, l'ús del sistema no va generar reduccions significatives en els delictes contra la propietat. D'altra banda, en

7. L'eslogan de Predpol era "More Than A Hotspot Tool" (Benbouzid, 2018).

8. Veure: <https://themarkup.org/prediction-bias/2023/10/02/predictive-policing-software-terrible-at-predicting-crimes>

una altra investigació, Saunders, Hunt i Hollywood (42) mostren com l'ús de la SSL (Strategic Subjects List) per part de la policia de Chicago, un sistema *person-based*, no es va demostrar eficaç en la prevenció de la violència relacionada amb l'ús d'armes de foc, ni a l'hora d'identificar potencials víctimes ni potencials autors.

El fet que hi hagi estudis que assenyalen que potser les eines predictives no compleixen amb les seves promeses no hauria de ser estrany si es fa una passa enrere. Per fer prediccions cal partir d'una sèrie de condicions inicials, d'una teoria sobre el fenomen que s'està analitzant (Varona Gómez, 2024), en aquest cas, la delinqüència. Segons Bennet Mosses i Chan (2018) aquestes assumpcions són que les dades disponibles reflecteixen la realitat de la delinqüència, que la delinqüència futura serà semblant a la delinqüència passada i que les variables omeses en els models són irrellevants.

Aquestes assumpcions no són realistes. En primer lloc, les dades que s'utilitzen normalment per "entrenar" els models són dades de delictes registrats per la policia. La qualitat d'aquestes dades està amenaçada per la xifra obscura: tots aquells delictes que per la no denúncia de les víctimes o per altres raons no arriben al coneixement de la policia i, per tant, no queden registrats (per exemple, Biderman i Reiss Jr., 1967). A més d'altres problemes d'inconsistència en els registres (Bennet Mosses i Chan, 2018). Les dades registrades per la policia reflecteixen abans el comportament de les organitzacions policials que el de la realitat delinqüencial (Becker, 2017). El problema és assumir que la policia és el consumidor final del producte quan és l'organització i les seves decisions (sobre quins delictes perseguir, per exemple) les que generen aquestes dades (Joh, 2017). Això genera bucles de retroalimentació (Ensing et al., 2018) o, dit d'una altra manera: la policia predictiva prediu l'activitat policial del futur, no la criminalitat futura (Lum i Isaac, 2016, p. 16). En segon lloc, assumir que la delinqüència futura serà igual o semblant a la del passat és problemàtic (Bennet Mosses i Chan, 2018), tal com va demostrar David Marsan (Benbouzid, 2018). Finalment, tot sistema predictiu es basa en l'assumpció que les variables no incloses en el model són irrellevants. Però aquesta és una assumpció obligada per la limitació d'informació i no basada en els fonaments de la teoria criminològica que, des dels seus diferents vessants, apunta en la gran complexitat del feno-

men i, per tant, en la gran quantitat de variables relacionades (p. e., Wikström i Kroneberg, 2022).

2.3.2. MAJOR EFICIÈNCIA

Una altra de les promeses que acompanya la implementació d'aquests sistemes és que la major precisió en les prediccions permetria una millor distribució dels recursos policials i, per tant, una major eficiència en la seva feina. A part que es pot posar en dubte l'afirmació sobre la major precisió amb el que s'ha vist en l'apartat anterior, aquesta promesa es basa en la idea que les organitzacions policials actuaran en funció de les alertes i les prediccions dels sistemes basats en IA.

En el cas dels sistemes predictius, que són en els que més s'ha estudiat aquest aspecte, la recerca ha destacat que el procés per passar de les prediccions a la seva aplicació pràctica pot ser complicat. Molt depèn de la manera com es transmet aquesta informació als agents de patrulla i de com s'ajusta a les prioritats ja establertes per la policia (Brayne i Christin, 2021; Ratcliffe, Tylor i Fisher, 2020). De fet, i en relació amb aquest últim punt, Selten, Robeer i Grimmelikhuijsen (2023) mostren com els agents de policia estan més inclinats a "fer cas" de les prediccions quan confirmen les seves opinions personals. Així, prediccions precises, per si soles, no garanteixen una millora en el rendiment policial.

2.3.3. MÉS JUSTÍCIA

Finalment, la promesa que la implementació d'aquests sistemes d'anàlisi, reconeixement i/o predicció portarà a decisions i actuacions més justes es basa en la idea que, en el cas de la policia, la seva utilització portarà a una reducció de la discrecionalitat dels agents i que el procés de processament i anàlisi de dades no discrimina de forma injusta (Bennet Mosses i Chan, 2018).

Pel que fa als sistemes de reconeixement facial, la investigació de Fussey et al. (2021) mostra que la discrecionalitat de l'organització i dels agents de policia continua sent molt significativa. Això es fa especialment evident a l'hora de determinar quins són els llocs on s'instal·len les càmeres de vigilància amb reconeixement facial i quines persones seran objecte d'observació. I sobre la no discriminació en el processament i anàlisi de les dades la recerca ha mostrat com aquests siste-

mes tenen molts problemes a l'hora de reconèixer persones racialitzades, especialment les dones, per una sobrerrepresentació a les dades d'entrenament d'imatges d'homes blancs (Buolamwini i Gebru, 2018). Això ha portat a problemes greus com la detenció de persones innocents⁹. De fet, aquests problemes són els que han portat a la prohibició general del seu ús a la Unió Europea que s'ha esmentat abans, a la prohibició en algunes ciutats dels Estats Units com San Francisco¹⁰ i al fet que algunes empreses deixin de vendre-li aquesta tecnologia a la policia¹¹.

En el cas dels sistemes predictius, aquesta assumpció és molt problemàtica per diferents aspectes. En primer lloc, aquests sistemes s'alimenten de dades de la mateixa activitat policial. Per tant, si l'activitat policial és esbiaixada d'origen, les prediccions també ho seran pel problema dels bucles de retroalimentació (Ensing et al., 2018; Lum i Isaac, 2016). La recerca criminològica¹² ha demostrat que l'activitat policial és selectiva amb determinats perfils de persones, principalment amb persones racialitzades o pertanyents a minories ètnico-racials¹³ (per exemple, Becket, Nyrop, Pflingst i Bowen, 2005; Bowling i Phillips, 2007; Majmundar i Weisburd, 2018; Medina Ariza, 2014). Una recerca periodística va posar-ho de manifest amb el mencionat software de Geolítica¹⁴. De fet, aquest és un aspecte destacat per Richardson, Schultz i Crawford (2019), el qual anomenen el problema de les dades contaminades: aquelles dades que han estat distorsionades per pràctiques com els abusos policials o biaixos en l'actuació policial (per exemple, actuacions basades en perfilatge ètnic).

Aquest problema és reconegut pels mateixos desenvolupadors dels sistemes, que han fet algunes propostes per tractar de reduir el seu impacte (Mohler et al., 2018), i assenyalen com a element crític la integritat de les dades (Brantingham, Vala-

sik i Mohler, 2018). Richardson et al. (2019), però, apunten que no hi ha procediments estandarditzats per a la recopilació, avaluació i ús de la informació obtinguda durant les activitats policials. A més, no existeixen evidències que les empreses que desenvolupen aquestes tecnologies verifiquin de manera independent la validesa de les dades policials. Les autores indiquen que és important tenir en compte que els biaixos en el treball policial poden manifestar-se de manera activa (quan són el resultat de pràctiques esbiaixades proactives) o bé a través de l'omissió d'informació sobre determinats delictes o infractors als quals no se'ls dedica la mateixa atenció que a d'altres. També és important entendre que els criteris de diferenciació que el model "aprèn" poden ser el resultat de processos de discriminació social històrics, de manera que els models poden discriminar per característiques físiques o de gènere, fins i tot si aquestes variables no estan presents en les dades d'entrenament de forma explícita (Adensamer i Klausner, 2021; Bennet Moses i Chan, 2018). Per exemple, donat que els processos de racialització als Estats Units estan molt relacionats amb la ubicació geogràfica i l'estatus socioeconòmic, és impossible destriar aquest efecte indirecte en les dades (Saphiro, 2017).

Davant aquest problema, es pot argumentar que els humans també prenen decisions esbiaixades i, per tant, en comparació, no seria un problema tan greu. Aquí hi hauria dos aspectes a tenir en compte sobre el perill de passar del biaix humà al biaix automatitzat. En primer lloc, la causa del segon és el primer, però a diferència dels humans, el fet que el segon estigui envoltat per un aura d'objectivitat, tecnologia i ciència el fa més difícil de reconèixer. En segon lloc, les decisions/recomanacions de sistemes automatitzats tenen el potencial d'afectar moltíssimes més persones que les decisions/recomanacions d'una persona o grups de persones amb prejudicis¹⁵ (Adensamer i Klausner, 2021).

9. Veure: <https://elpais.com/tecnologia/2023-08-12/la-detencion-erronea-de-una-embarazada-reaviva-el-rechazo-a-los-sistemas-de-reconocimiento-facial.html>

10. Veure: https://elpais.com/tecnologia/2019/05/15/actualidad/1557904606_766075.html

11. Veure: <https://www.bbc.com/news/business-52989128>

12. De fet ha sigut reconegut per alguns comandaments policials. Veure, per exemple, <https://www.3cat.cat/324/eduard-sallent-cap-dels-mossos-admet-biaix-etnic-en-les-identificacions-policials/noticia/3021431/>

13. El professor Mike Shor va fer un recull de més de 900 estudis que mostren el biaix racial a les pràctiques policials: <https://policeresearch.us/studies/>

14. Veure: <https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them>

15. De fet aquest és un dels principals perills assenyalats per Cathy O'Neil (2016) al seu llibre *Weapons of math destruction*.

3. Tecnologia, ciència i el debat públic

3.1. TECNOLOGIA I CIÈNCIA

A vegades, l'ús de la tecnologia sembla conferir una mena de legitimitat científica a pràctiques que, en realitat, no tenen una base sòlida de fonaments científics, o fins i tot n'estan completament desproveïdes. Aquesta confiança en la tecnologia pot crear una percepció errònia que les decisions són més objectives o fiables quan, de fet, poden estar basades en dades limitades, suposicions no verificades o fins i tot errors en la seva aplicació. La tecnologia, lluny de ser una garantia de rigor, pot acabar essent una disfressa que amaga les deficiències metodològiques o les incongruències d'algunes pràctiques.

Per exemple, quan s'introdueix una nova tecnologia, sempre va acompanyada d'una sensació d'evolució o progrés. Pel que fa als sistemes predictius que utilitzen algorismes d'aprenentatge automàtic, es presenten amb la promesa de ser més precisos que les tècniques predictives anteriors. La idea és que, a mesura que els models es fan més complexos, la seva capacitat per predir resulta més acurada. Tanmateix, és important destacar que, segons Hand (2006), l'augment de la complexitat dels models predictius no suposa una millora substancial en termes de precisió respecte als models més senzills. De fet, l'estudi de Hand (2006) demostra que el valor afegit dels models més complexos és relativament petit en comparació amb els models més simples, pel que fa a la seva capacitat per fer prediccions exactes. A més, l'autor també argumenta que, tot i les promeses de millora, els models complexos tenen algunes limitacions importants. Aquests models solen requerir una gran quantitat d'assumpcions sobre les dades i el seu comportament, moltes de les quals no es compleixen en la pràctica. En altres paraules, els models més sofisticats sovint es basen en condicions teòriques que no es donen al món real, fet que pot comprometre la seva eficàcia. En canvi, els models més senzills, tot i que menys complexos, tenen el gran avantatge de ser més transparents i fàcilment interpretables. Això significa que és més senzill entendre com funcionen i per què arriben a determinats resultats, la

qual cosa és fonamental, especialment en àmbits on la comprensió i la justificació dels resultats són essencials (com en el cas de la seguretat ciutadana o la justícia penal). Per tant, tot i que els models complexos poden semblar atractius per la seva capacitat de processar grans quantitats de dades, sovint els models més senzills ofereixen una millor combinació de precisió i transparència.

La “il·lusió de progrés” de la que parlava Hand (2006) ha estat confirmada en estudis posteriors. Un exemple rellevant el trobem en una investigació liderada per Salganik et al. (2006), en la qual es va dur a terme una col·laboració entre una gran varietat d'investigadors i investigadores de diferents disciplines. L'objectiu principal d'aquesta investigació era, utilitzant un conjunt de dades comú molt ric en variables, predir sis fites en les trajectòries vitals de les persones. A partir d'aquest conjunt de dades, diversos grups d'investigadors van desenvolupar els seus propis models predictius per intentar anticipar aquests esdeveniments futurs. Els resultats d'aquest estudi van mostrar una troballa sorprenent: els models més complexos, aquells que utilitzaven aprenentatge automàtic, no van demostrar ser més precisos que els models més senzills. Però l'ensenyament més important d'aquest estudi és que, fins i tot els millors models, aquells que van oferir les prediccions més exactes, no van ser realment molt precisos.

Aquest fet suggereix que la capacitat de predir determinats esdeveniments futurs, fins i tot utilitzant tècniques avançades, té límits estrictes. En una presentació¹⁶, Narayanan, un dels participants en l'estudi esmentat, va expressar que, tot i que els algorismes per detectar patrons de discurs a les xarxes socials poden tenir un gran potencial, en el cas de la policia predictiva, la implementació de la IA planteja molts problemes. Segons ell, “aquests problemes són difícils perquè no podem predir el futur. Això hauria de ser sentit comú, però sembla que hem decidit suspendre el sentit comú quan la intel·ligència artificial està involucrada”. Aquesta reflexió posa de manifest un aspecte fonamental: l'ús de la IA en camps tan sensibles com la seguretat pública comporta riscos importants, ja que es fonamenta

16. Disponible a: <https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf>

en la suposició que la tecnologia pot oferir respostes més fiables, sense reconèixer les limitacions inherents a qualsevol sistema predictiu que intenti anticipar comportaments humans en contextos tan complexos i imprevisibles com aquest.

També hi ha casos que directament no tenen fonaments científics. En el cas dels sistemes de reconeixement de trets biomètrics, alguns algorismes prometen detectar emocions, mentides o la propensió a la criminalitat a partir d'imatges. Per exemple, grans empreses tecnològiques, com Google i IBM, venen sistemes de reconeixement d'emocions mitjançant expressions facials, tot i que no hi ha evidència científica que permeti inferir les emocions d'una persona d'aquesta manera (Barrett et al., 2020). Altres tecnologies asseguruen poder detectar mentides. Un exemple és el sistema iBorderCtrl, que utilitza models d'aprenentatge automàtic per identificar comportaments sospitosos (Sánchez-Monedero i Dencik, 2020). De fet, des del polígraf és conegut que hi ha poc suport científic per a la idea que les persones generen una resposta física quan expliquen un fet que saben que és fals (ja que, per exemple, aquestes mateixes respostes físiques poden aparèixer independentment de si el contingut del relat és fals o verídic) o que la mentida sigui un indicador vàlid de culpabilitat (National Research Council, 2003; Sánchez-Monedero i Dencik, 2020).

El cas citat de VeriPol, el sistema utilitzat pel Cos Nacional de Policia, podria ser diferent, ja que funciona a partir de processament del llenguatge natural (Quijano Sánchez et al., 2019). A falta d'avaluacions independents, principalment per la manca de transparència (Martínez Garay et al., 2022), però, es podria dir que presenta limitacions com la representativitat de les dades d'entrenament (Martínez Garay et al., 2022) i el fet que no detecta mentides, sinó similituds amb denúncies prèviament catalogades com a falses per la mateixa policia¹⁷. Alguns investigadors van molt més enllà i afirmen que tenen models que poden predir la "tendència criminal" d'una persona a partir d'imatges d'aquestes (p. e. Hashemi i Hall, 2020). Tanmateix, aquests estudis han estat criticats per la seva falta de rigor científic i per perpetuar prejudicis racials (Bowyer et al., 2020).

Finalment, s'ha de tenir en compte que els objectius dels desenvolupadors dels sistemes de reconeixement o predictius basats en IA són abans comercials que científics. Tot i que l'ús de la tecnologia sembla embolcallar certes pràctiques amb una capa d'aparença científica (Hong, 2023; Jefferson, 2018), les persones que desenvolupen aquesta tecnologia no sempre segueixen els mateixos procediments ni tenen com a objectiu principal ampliar la base del coneixement científic. Un exemple clar d'això es troba en l'estudi de Benbouzid (2018) que revela les diferències entre el sismòleg Marsan i els desenvolupadors de Geolítica en aquest aspecte. Per Marsan, la manera d'avaluar l'èxit del seu model predictiu consistia a comparar el model matemàtic amb una visió coherent del fenomen que s'estava estudiant, buscant una aproximació que reflectís fidelment la realitat que volia modelitzar. L'objectiu últim era entendre millor els terratrèmols. En canvi, els creadors del sistema de Geolítica valoraven l'èxit del seu model d'una manera diferent: el mesuraven a través de la seva capacitat de superar altres models d'altres empreses en termes de rendiment. Aquesta distinció revela una diferència fonamental en les seves aproximacions, ja que mentre que Marsan prioritzava una comprensió profunda de l'objecte estudiat, els desenvolupadors del sistema predictiu PredPol estaven més enfocats en la millora competitiva del seu model respecte a altres alternatives tecnològiques.

3.2. EL DEBAT PÚBLIC

El debat públic sobre aquestes eines es fa difícil per dues raons principalment. En primer lloc, pels problemes de transparència i rendició de comptes. En segon lloc, a conseqüència de la tecnificació del debat.

Aquests *softwares* basats en IA presenten dues grans dificultats pel que fa a la transparència. D'una banda, molts d'aquests algorismes estan protegits per drets de propietat intel·lectual, com el secret comercial. S'ha vist que fins i tot quan les eines no estan protegides per secret comercial, ja que, per exemple, són propietat d'administracions públiques, també hi ha una manca de transparència sobre el seu funcionament, com és el cas d'algunes que s'utilitzen a l'Estat espanyol

17. Veure: <https://elpais.com/tecnologia/2021-03-08/veripol-el-poligrafo-inteligente-de-la-policia-puesto-en-cuestion-por-expertos-en-etica-de-los-algoritmos.html>

(Martínez Garay et al., 2022). De l'altra, la seva naturalesa de "caixa negra" dificulta enormement la comprensió del seu funcionament, fins i tot per aquells que els han dissenyat (Benbouzid, 2018; Bennet Moses i Chan, 2018; Ferguson, 2017; Joh, 2017; Saphiro, 2017). Poc sovint les empreses proporcionen informació clara sobre com operen els seus sistemes, quins tipus de dades es fan servir o quines estratègies es duen a terme per garantir la rendició de comptes davant possibles perills (Martínez Garay et al., 2022; Richardson et al., 2019).

Aquest problema de transparència genera també un altre repte: la rendició de comptes. En una democràcia, es pressuposa que les accions de les autoritats i institucions públiques han de ser comprensibles per la ciutadania. Però en el cas de l'ús d'eines com aquestes, les persones que segueixen les recomanacions dels sistemes sovint no entenen com funcionen, fet que complica la justificació de les decisions preses. A més, l'ús de tècniques de policia predictiva suposa un altre obstacle en aquest sentit, ja que genera un buit de responsabilitat. És complicat assignar responsabilitats quan intervenen múltiples actors, com els desenvolupadors, els responsables de la presa de decisions i els operadors dels sistemes, i resulta difícil establir clarament qui és responsable de què (Bennet Moses i Chan, 2018). Com apunten Adensamer i Klausner (2021), la decisió d'introduir aquests sistemes recau sovint en les persones amb més poder dins l'organització, o en els responsables polítics, i com que molt probablement no entenen com funcionen aquestes eines, resulta difícil demanar comptes als agents de policia implicats. Segons Ferguson (2017), aquesta dificultat s'agreuja per la ja existent manca de transparència i rendició de comptes dins els cossos policials.

En segon lloc, una tasca fonamental és entendre que l'ús d'eines basades en IA en el treball policial no es redueix simplement a una qüestió tècnica. Cal ser conscient de què està en joc quan es debat sobre qualsevol aspecte relacionat amb la seguretat o les funcions policials (Loader, 2020). Un dels aspectes que sovint es perd de vista en aquests debats centrats en els

aspectes tècnics (algorismes, models predictius, precisió, eficiència, etc.) és la finalitat de l'ús d'aquestes eines: per a quins problemes, per a quins delictes, quines persones, quins llocs i, en definitiva, quins riscos s'intentarà identificar o predir (Hong, 2023; Vestby i Vestby, 2021). Per exemple, en el debat públic s'hauria de justificar perquè s'han desenvolupat més eines per predir delictes com furtos o robatoris en domicilis que per predir delictes com frau, delictes fiscals, blanqueig, etc.¹⁸. També és cridaner que, des de 2016, existeixen eines per predir males pràctiques policials però no s'utilitzen¹⁹. A més, hi ha altres qüestions que són més normatives que tècniques, com ara quins objectius es volen assolir en utilitzar aquestes eines (Vestby i Vestby, 2021) o què hauria de prevaldre en els sistemes, si la precisió o l'equitat (Berk, 2021).

D'altra banda, cal abordar el problema de les dades contaminades, que tot i ser reconegut, resulta difícil de gestionar per dues raons principals: hi ha poc control sobre les pràctiques policials problemàtiques i no existeix cap solució tècnica que pugui mitigar el problema de la informació no generada per l'omissió de fets (Richardson et al., 2019). Addicionalment, Wang et al. (2024) apunten altres problemes de legitimitat relacionats amb el que anomenen "optimització predictiva" (l'ús de l'aprenentatge automàtic per predir comportaments individuals i prendre decisions basades en aquestes prediccions). Alguns d'aquests problemes inclouen: si els objectius de l'optimització predictiva són realment socialment rellevants, si les persones afectades per aquestes decisions tenen la capacitat de qüestionar-les, la gravetat de les conseqüències d'un error de classificació, si les persones participen de manera voluntària o involuntària en els processos d'optimització predictiva, i si les persones amb més coneixement del sistema poden sortir-ne beneficiades.

18. En aquest sentit, Clifton, Lavigne i Tseng de *The New Inquiry Magazine* van desenvolupar una eina predictiva amb una clara intenció reivindicativa: <https://whitecollar.thenewinquiry.com/>

19. Veure <https://fivethirtyeight.com/features/we-now-have-algorithms-to-predict-police-misconduct/>

4. En conclusió: és la IA una nova revolució en la policia? Hauria de ser-ho?

Tot i que la IA genera un gran interès i s'ha presentat en molts discursos com una “nova revolució” tecnològica, la seva incorporació al treball policial no sembla que comporti un canvi tan profund com el que van suposar altres innovacions tecnològiques al llarg de la història, com el cotxe, la ràdio o el telèfon. Aquestes últimes van transformar la naturalesa del treball policial. La IA, en canvi, sembla estar més enfocada a l'automatització de tasques específiques. Això pot ser útil per a certes funcions, però no sembla que canviï el paradigma de la feina policial de manera fonamental.

En aquest sentit, potser el problema principal és el rebombori mediàtic generat entorn de la IA. Aquesta exaltació excessiva pot eclipsar els riscos i problemes inherents al seu ús. La promesa de la IA pot conduir a l'adopció de solucions tecnològiques sense una reflexió prèvia sobre els seus fonaments

científics i les seves implicacions ètiques i socials. De fet, ja s'ha comprovat que la IA pot acabar replicant i accentuant problemes existents de l'activitat policial, com ara els prejudicis socials o l'ús inapropiat d'informació. A mesura que els models d'IA s'entrenen amb dades històriques, hi ha el risc que perpetuin els mateixos biaixos que ja existeixen, posant en perill la igualtat de tracte i la justícia.

Finalment, el debat sobre la IA a la policia sovint es concentra en els aspectes tècnics i els beneficis immediats, però potser cal centrar el debat, per exemple, en les llums i ombres del paper preventiu de la policia o, de forma més general, en les funcions de la policia en societats democràtiques (per exemple, Loader, 2020). Sense aquestes reflexions prèvies, l'adopció de la IA podria acabar convertint-se en un simulacre de solució.



Referències

- Adensamer, F., & Klausner, A. (2021). "Part Man, Part Machine, All Cop": Automation in Policing. *Frontiers in Artificial Intelligence*, 4, 1-10. <https://doi.org/10.3389/frai.2021.655486>
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1-68. <https://doi.org/10.1177/1529100619832930>
- Becker, H. S. (2017). *Evidence*. University of Chicago Press.
- Beckett, K., Nyrop, K., Pfingst, L., & Bowen, M. (2005). Drug use, drug possession arrests, and the question of race: Lessons from Seattle. *Social Problems*, 52(3), 419-441. <https://doi.org/10.1525/sp.2005.52.3.419>
- Benbouzid, B. (2018). Values and consequences in predictive machine evaluation: A sociology of predictive policing. *Science & Technology Studies*, 32(4), 119-136.
- Benbouzid, B. (2019). To predict and to manage. Predictive policing in the United States. *Big Data & Society*, 6(1). <https://doi.org/10.1177/2053951719861703>
- Bennett Moses, L., & Chan, J. (2018). Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing & Society*, 28(7), 806-822. <https://doi.org/10.1080/10439463.2016.1253695>
- Berk, R. A. (2021). Artificial intelligence, predictive policing, and risk assessment for law enforcement. *Annual Review of Criminology*, 4, 209-237. <https://doi.org/10.1146/annurev-criminol-051520-012342>
- Biderman, A. D., & Reiss Jr, A. J. (1967). On exploring the "dark figure" of crime. *The Annals of the American Academy of Political and Social Science*, 374(1), 1-15. <https://doi.org/10.1177/000271626737400102>
- Bowling, B., & Phillips, C. (2007). Disproportionate and discriminatory: Reviewing the evidence on police stop and search. *The modern law review*, 70(6), 936-961. <https://doi.org/10.1111/j.1468-2230.2007.00671.x>
- Bowyer, K. W., King, M. C., Scheirer, W. J., & Vangara, K. (2020). The "criminality from face" illusion. *IEEE Transactions on Technology and Society*, 1(4), 175-183.
- Brantingham, P. (2017). The logic of data bias and its impact on place-based predictive policing. *Ohio State Journal of Criminal Law*, 15, 473-486.
- Brantingham, J., Valasik, M., & Mohler, G. (2018). Does predictive policing lead to biased arrests? Results from a randomized controlled trial. *Statistics and Public Policy*, 5(1), 1-6. <https://doi.org/10.1080/2330443X.2018.1438940>
- Brayne, S., & Christin, A. (2021). Technologies of crime prediction: The reception of algorithms in policing and criminal courts. *Social Problems*, 68(3), 608-624. <https://doi.org/10.1093/socpro/spaa004>
- Brayne, S., Rosenblat, A. & Boyd, D. (2015). *Predictive policing*. Disponible a http://www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77-91. Disponible a <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>
- Calatayud, A. (2021). *Report Automating Society 2020*. Disponible a <https://automatingsociety.algorithwatch.org/report2020/spain>
- Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C., & Venkatasubramanian, S. (2018). Runaway feedback loops in predictive policing. *Proceedings of Machine Learning Research*, 81, 1-12. Disponible a <http://proceedings.mlr.press/v81/ensign18a/ensign18a.pdf>
- Ferguson, A. G. (2017). Policing predictive policing. *Washington University Law Review*, 94(5), 1109-1189.

Fussey, P., Davies, B., & Innes, M. (2021). 'Assisted' facial recognition and the reinvention of suspicion and discretion in digital policing. *The British Journal of Criminology*, 61(2), 325-344.

Goold, B. J. (2004). *CCTV and policing: Public area surveillance and police practices in Britain*. Oxford University Press.

Hand, D. J. (2006). Classifier technology and the illusion of progress. *Statistical Science*, 21(1), 1-14. <https://projecteuclid.org/journals/statistical-science/volume-21/issue-1/Classifier-Technology-and-the-Illusion-of-Progress/10.1214/0883423060000000060.full>

Harcourt, B. E. (2007). *Against prediction: profiling, policing and punishment in an actuarial age*. The University of Chicago Press.

Hashemi, M., & Hall, M. (2020). Retracted article: Criminal tendency detection from facial images and the gender bias effect. *Journals of Big Data*, 7(2), 1-16. <https://doi.org/10.1186/s40537-019-0282-4>

Hong, S.-h. (2023). Prediction as extraction of discretion. *Big Data & Society*, 10(1), 1-11. <https://doi.org/10.1177/20539517231171053>

Hunt, P., Saunders, J., & Hollywood, J. S. (2014). *Evaluation of the Shreveport Predictive Policing Experiment*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR531.html

Jansen, F. (2018, May 7). *Data driven policing in the context of Europe* [Working paper]. DATA-JUSTICE, Cardiff University. <https://datajustice-project.net>

Jansen, F., Sánchez-Monedero, J., & Dencik, L. (2021). Biometric identity systems in law enforcement and the politics of (voice) recognition: The case of SiiP. *Big Data & Society*, 8(2), 1-13. <https://doi.org/10.1177/20539517211063604>

Jefferson, B. J. (2018). Predictable policing: Predictive crime mapping and geographies of policing and race. *Annals of the American Association of Geographers*, 108(1), 1-16. <https://doi.org/10.1080/24694452.2017.1293500>

Joh, E. E. (2017). Feeding the Machine: Policing, Crime Data, & Algorithms. *William & Mary Bill of Rights Journal*, 26(2), 287-302.

Loader, I. (2020). *Revisiting the police mission*. Policing Review. Disponible a https://www.policingreview.org.uk/wp-content/uploads/insight_paper_2.pdf

Lum, K. i Isaac, W. (2016). To Predict and Serve? *Significance*, 13(5), 14-19. <https://doi.org/10.1111/j.1740-9713.2016.00960.x>

MacAvaney, S., Yao, H. R., Yang, E., Russell, K., Goharian, N., & Frieder, O. (2019). Hate speech detection: Challenges and solutions. *PLoS one*, 14(8), e0221152. <https://doi.org/10.1371/journal.pone.0221152>

Majmundar, M. K., & Weisburd, D. (Eds.). (2018). *Proactive policing: Effects on crime and communities*. National Academies Press.

Manning, P. K. (2008). *The Technology of Policing: Crime Mapping, Information Technology, and the Rationality of Crime Control*. New York University Press.

Martínez Garay, L. (coord.), Boix Palop, A., Briz Redón, A., Flores Giménez, F., García Ortiz, A., Molina Sánchez, M., ... Soriano Aranz, A. (2022). *Three predictive policing approaches in Spain: Viogén, RisCanvi and Veripol: Assessment from a human rights perspective*. Disponible a <https://regulation.blogs.uv.es/files/2024/09/Three-predictive-policing-approaches-web.pdf>

Medina Ariza, J. J. (2014). Police-initiated contacts: young people, ethnicity, and the "usual suspects". *Policing and Society*, 24(2), 208-223. <https://doi.org/10.1080/10439463.2013.784301>

Miralles, I., Campisi, L., & Díaz, P. (2021). *Vigilancia hi-tech en tiempos del COVID-19*. Disponible a <https://www.donestech.net/files/mass-surveillance-esp-cast.pdf>

Miró Llinares, F. (2020). Predictive policing: Utopia or dystopia? On attitudes towards the use of big data algorithms for law enforcement. *Revista d'Internet, Dret i Política*, (30), 1-18.

Mohler, G., Raje, R., Carter, J., Valasik, M., & Brantingham, J. (2018). A penalized likelihood method for balancing accuracy and fairness in predictive policing. 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), October 7-10, 2018. Disponible a <https://ieeexplore.ieee.org/abstract/document/8616417/authors#authors>

Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., & Brantingham, P. J. (2015). Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, 110(512), 1-43. <https://doi.org/10.1080/01621459.2015.1077710>

Moll, I. (2021). The Myth of the Fourth Industrial Revolution. *Theoria*, 68(167), 1-38. <https://doi.org/10.3167/th.2021.6816701>

Moll, I. (2023). Why there is no technological revolution, let alone a “Fourth Industrial Revolution”. *South African Journal of Science*, 119(1/2), 1-6. <https://doi.org/10.17159/sajs.2023/12916>

Moore, M.H. (2003), Sizing up CompStat: an important administrative innovation in policing. *Criminology & Public Policy*, 2(3), 469-494. <https://doi.org/10.1111/j.1745-9133.2003.tb00009.x>

Narayanan, A. i Kapoor, S. (2024). *AI snake oil: What artificial intelligence can do, what it can't, and how to tell the difference*. Princeton University Press.

National Research Council. (2003). *The polygraph and lie detection*. The National Academies Press. <https://doi.org/10.17226/10420>

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Purshouse, J., & Campbell, L. (2022). Automated facial recognition and policing: a Bridge too far? *Legal Studies*, 42(2), 209-227. <https://doi.org/10.1017/lst.2021.22>

Quijano Sánchez, L., Liberatore, F., Camacho Collados, J., & Camacho Collados, M. (2018). Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police. *Knowledge-Based Systems*, 149, 155-168. <https://doi.org/10.1016/j.knosys.2018.02.017>

Ratcliffe, J. H., Taylor, R. B., & Fisher, R. (2020). Conflicts and congruencies between predictive policing and the patrol officer's craft. *Policing & Society*, 30(6), 639-655. <https://doi.org/10.1080/10439463.2019.1577844>

Ratcliffe, J. H., Taylor, R. B., Askey, A. P., Thomas, K., Grasso, J., Bethel, K. J., Fisher, R., & Koehnlein, J. (2021). The Philadelphia predictive policing experiment. *Journal of Experimental Criminology*, 17, 15-41. <https://doi.org/10.1007/s11292-019-09400-2>

Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *New York University Law Review Online*, 94(15), 15-55.

Sabaté Delgado, J. (1984). La profesión policial en España: un análisis sociológico. *Papers: Revista de Sociologia*, (21), 133-152.

Salganik, M. J., Lundberg, I., Kindel, A. T., Ahearn, C. E., Al-Ghoneim, K., Almaatouq, A., ... & McLanahan, S. (2020). Measuring the predictability of life outcomes with a scientific mass collaboration. *Proceedings of the National Academy of Sciences*, 117(15), 8398-8403. <https://doi.org/10.1073/pnas.1915006117>

Sánchez-Monedero, J., & Dencik, L. (2022). The politics of deceptive borders: “Biomarkers of deceit” and the case of iBorderCtrl. *Information, Communication & Society*, 25(3), 413-430. <https://doi.org/10.1080/1369118X.2021.1912628>

San Martín Segura, D. (2020). El papel de los objetos técnicos en el giro preventivo de las políticas securitarias. *Revista Crítica Penal y Poder*, (19), 30-55.

Saunders, J., Hunt, P., & Hollywood, J. S. (2016). Predictions put into practice. *Journal of Experimental Criminology*, 12, 347-371. <https://doi.org/10.1007/s11292-016-9272-0>

Selten, C., Robeer, M., & Grimmelikhuijsen, S. (2023). “Just like I thought”: Street-level bureaucrats trust AI recommendations if they confirm their professional judgment. *Public Administration Review*, 83(2), 263-278. <https://doi.org/10.1111/puar.13602>

Shapiro, R. (2017). Reform predictive policing. *Nature*, 541(7638), 458-460. <https://doi.org/10.1038/541458a>

Shapiro, A. (2019). Predictive policing for reform? Indeterminacy and intervention in big data policing. *Surveillance & Society*, 17(3/4), 456-472. <https://doi.org/10.24908/ss.v17i3/4.10410>

Stroshine, M. S. (2015). Technological innovations in Policing. A R. G. Dunham i G. P. Alpert (eds.), *Critical issues in Policing: Contemporary Readings* (7th Edition) (pp. 229-243). Waveland Press.

Varona Gómez, D. (2024). Algoritmos e inteligencia artificial en el sistema de justicia penal (editorial). InDret: *Revista para el Análisis del Derecho*, (4), 1-4. Disponible a <https://indret.com/wp-content/uploads/2024/10/1893.pdf>

Vestby, A., & Vestby, J. (2021). Machine learning and the police: Asking the right questions. *Policing*, 15(1), 44-58. <https://doi.org/10.1093/policing/pz035>

Wang, A., Kapoor, S., Barocas, S., & Narayanan, A. (2024). Against predictive optimization: On the legitimacy of decision-making algorithms that optimize predictive accuracy. *ACM Journal on Responsible Computing*, 1(1), 1-45. <https://doi.org/10.1145/3636509>