Artificiell intelligens inom brottsbekämpning

En systematisk litteraturgenomgång av algoritmer och AI inom brottsbekämpande arbete

Kandidatarbete inom Industriell ekonomi

LUDVIG DIGNÉ
GABRIEL GERAGHTY
JAKOB HANÁK
ADAM LANDBERG
AGNES ROSENDAL
CARL D. TORBJÖRNSSON

INSTITUTIONEN FÖR TEKNIKENS EKONOMI OCH ORGANISATION
AVDELNINGEN FÖR TEKNIK, VETENSKAP OCH SAMHÄLLE
CHALMERS TEKNISKA HÖGSKOLA
Göteborg, Sverige 2022
www.chalmers.se

Kandidatarbete TEKX04-22-21

Kandidatarbete TEKX04-22-21

Artificiell intelligens inom brottsbekämpning

En systematisk litteraturgenomgång av algoritmer och AI inom brottsbekämpande arbete

Artificial intelligence in law enforcement

A systematic literature review on the use of algorithms and AI for crime-fighting

LUDVIG DIGNÉ
JAKOB HANÁK
AGNES ROSENDAL
GABRIEL GERAGHTY
ADAM LANDBERG
CARL D. TORBJÖRNSSON

TEKNIKENS EKONOMI OCH ORGANISATION
Avdelning för teknik, vetenskap och samhälle
CHALMERS TEKNISKA HÖGSKOLA
Göteborg, Sverige 2022

Artificiell intelligens inom brottsbekämpning
En systematisk litteraturgenomgång av algoritmer och AI inom brottsbekämpande arbete

LUDVIG DIGNÉ
JAKOB HANÁK
AGNES ROSENDAL
GABRIEL GERAGHTY
ADAM LANDBERG
CARL D. TORBJÖRNSSON

Kandidatarbete TEKX04-22-21
Teknikens ekonomi och organisation
Chalmers tekniska högskola
412 96 Göteborg
Sverige
Telefon + 46 (0)31-772 1000

Göteborg, Sverige 2022
Gothenburg, Sweden 2022

Artificial intelligence in law enforcement
A systematic literature review on the use of algorithms and AI for crime-fighting

LUDVIG DIGNÉ GABRIEL GERAGHTY
JAKOB HANÁK ADAM LANDBERG
AGNES ROSENDAL CARL D. TORBJÖRNSSON

Department of Technology Management and Economics
Chalmers University of Technology

SUMMARY
AI research is a rapidly growing field, and technology that uses AI is entering more and more parts of society, including law enforcement. However, the available literature in the field has yet to be compiled together, something which this report seeks to remedy. In this study, we summarize and synthesize the claims and findings from previous literature, in particular relating to the possibilities and risks associated with using algorithms and artificial intelligence in police work. The method of study is a systematic literature review, using results from searches in the Scopus academic research database. The keywords used are "law enforcement", "policing", "crime prevention", "crime reduction", and "surveillance", in combination with "algorithm" or "artificial intelligence". The results show that the use of AI has great potential to facilitate law enforcement, enabling the use of facial recognition, crime predictions, tracking of individuals, etc. Most of the described use cases involve either the analysis of images, behavior or text, or a combination of these. At the same time, the technology also constitutes a number of ethical risks as it can lead to discrimination, ambiguity regarding where the responsibility for the outcome of the AI programs lies, mass surveillance, and in particular, the infringing of privacy. Despite these questions, we find that most of the literature either advocates the implementation of digital technology, or outlines frameworks for it. The positive aspects and the technological advances associated with the implementation of AI are typically emphasized, more so than the potential risks. We also find that the articles that touch on the same subjects are consistent when it comes to their descriptions of current capabilities, the potential of technology, and what topics future research should explore.

Keywords: Artificial intelligence, algorithms, law enforcement, policing, surveillance, crime prevention, ethics

Note: The report is written in Swedish

SAMMANFATTNING

Forskning kring AI är ett snabbt växande område, och teknik som använder sig av AI letar sig in i allt fler områden i samhället, däribland brottsbekämpningen. Det saknas dock en sammanställning av den tillgängliga litteraturen på området, vilket är något som denna rapport ämnar åstadkomma. I den här rapporten redogör och syntetiserar vi de gemensamma redogörelserna från tidigare studier om möjligheter och risker med att använda algoritmer och artificiell intelligens inom polisiärt arbete. Arbetets metod har utformats som en systematisk litteraturgenomgång av resultaten från sökningar i databasen Scopus, med sökorden "law enforcement", "policing", "crime prevention", "crime reduction", "surveillance" i kombination med "algorithm" eller "artificial intelligence". Resultaten visar att användningen av AI har stor potential att underlätta för brottsbekämpningen, med möjligheter som ansiktsigenkänning, brottsförutsägelser, spårning av individer m. fl.. Det går att bryta ner de allra flesta exemplen på användning av AI till bild-, beteende- och textanalys. Samtidigt utgör detta även ett flertal etiska risker, då det kan leda till diskriminering, oklarhet gällande var ansvaret för AI-programmets agerande ligger, massövervakning och inte minst kränkning av den personliga integriteten. Trots dessa frågor finner vi att merparten av litteraturen antingen förespråkar implementering av digital teknik, eller ramverk för digital teknik. Det fokuseras oftast huvudsakligen på de positiva aspekterna och de tekniska framstegen associerade med en implementering av AI, snarare än potentiella risker. Vi finner även att de artiklar som berör samma ämnen är tillika samstämmiga i frågor om dagens förmågor, teknikens potential, och vad de vill se framtida arbeten kring.

Nyckelord: Artificiell intelligens, algoritmer, brottsbekämpning, polisarbete, övervakning, brottsförebyggande, etik

Ordlista

Artificiell Intelligens (*AI*) - förmåga hos datorprogram att efterlikna människors kognitiva förmågor och intelligens.

Big Data - ett begrepp för den datamängd som växt fram på senare år, och syftar på datamängder av sådan storlek att de är svåra att bearbeta med traditionella metoder. (sv. stordata)

Computer Vision - förmågan hos dator att förstå och extrahera innehåll från bilder beroende på vad söks. (sv. datorseende)

Convolutional Neural Network (CNN) - en teknik som används vid bildanalys som har visat sig väldigt bra på att känna igen ansikten under svåra förhållanden. (sv. konvolutionella neurala nätverk)

Crowdsensing - en teknik där en stor grupp av individer med mobila enheter kapabla till avkänning och beräkning, kollektivt delar och utvinner information för ett gemensamt intresse.

Explainable AI (XAI) - AI vars logik går att förklara och är begriplig för människor. (sv. förklarbar AI)

Internet of Things (IoT) - nätverket av apparater utrustade med sensorer, mjukvara o.s.v. som använder internet för att kommunicera genom ett utbyte av olika former av data. (sv. sakernas internet)

Unsupervised Learning - inträffar när AI:n tar in data och upptäcker mönster utan att en människa är med och medverkar. (sv. oövervakad maskininlärning)

Innehållsförteckning

1. Inledning

På 1940-talet uppfanns den första datorn i modern bemärkelse, och det dröjde inte lång tid tills den unge akademikern Alan Turing började fundera kring beräkningsmaskinens framtida potential (Anyoha, 2017). Trots den mycket begränsade prestandan hos den dåvarande datorn ställde han sig frågan om inte datorer i framtiden skulle ha förmågan att tänka på ett sätt som levde upp till eller överträffade människan. Människan använder sig av information och reson för att lösa problem och ta beslut, så varför skulle inte maskiner kunna göra samma sak? Alan dog 1956, men två år senare plockade Marvin Minsky upp stafettpinnen och var tillsammans med John McCarthy värd för Dartmouth Summer Research Project on Artificial Intelligence. De bjöd in välrespekterade forskare från olika områden för en öppen diskussion kring *artificiell intelligens* (AI). Hädanefter blev artificiell intelligens en vedertagen term. Forskningen fortsatte och området utvecklades, men datorernas lagringskapacitet och beräkningskraft var begränsande faktorer, och intresset för AI svalnade (Anyoha, 2017). Under 80-talet såg AI ytterligare ett uppsving till följd av allt större forskningsanslag och i kombination med den snabba utvecklingen av prestanda under 90-talet kunde det konstateras att AI hade kommit för att stanna (Anyoha, 2017; Business Standard, 2017).

Idag lever vi i en värld där AI har letat sig in i allt fler samhällsområden och totalt förändrat hur vi jobbar. Politik, utbildning, sjukvård, näringslivet och juridik är bara några av de områden i samhället som genomgått stora förändringar till följd av detta. Ett område som blivit högaktuellt på senare tid är brottsbekämpning, där traditionella rutiner fortfarande dominerar. Majoriteten av polisarbetet som sker görs först efter att brottet har inträffat, vilket innebär att det krävs stora resurser i form av personal, pengar och tid. Dessutom minskar möjligheterna att klara upp brottet ju längre tid som går. Med hjälp av AI kan polisen utföra mer prediktivt polisarbete, det vill säga att förutsäga var, hur, och när brott kommer att begås för att stå bättre rustade för utryckning eller förhindrande av brott (Innefu, u.å.). Förhoppningen är att man med hjälp av AI kommer besitta förmågan att analysera material från videokameror i realtid. Ett annat prospekt är att kunna läsa av sociala medier för att identifiera mönster och beteenden som kan vara indikationer på att brott är på väg att ske. En lyckad implementering av detta skulle kunna rädda många människoliv (Freeman, 2020), spara enorma resurser, och minimera risken för mänskliga fel som kan uppstå vid dagens brottsutredning (Suralkar et al., 2020). Dessa är bara några av de områden där det föreslås att AI kan ge upphov till stora förbättringar.

Interpol (2020) menar i sin AI-rapport att användningen av AI inom rättsväsendet måste ske på ett sådant sätt att det tar hänsyn till de generella principerna för mänskliga rättigheter, demokrati, rättvisa, och den rådande lagstiftningen. För att lyckas med detta hänsynstagande måste myndigheter inom rättsväsendet arbeta för att uppnå kraven *fairness, accountability, transparency,* och *explainability* (sv. rättvisa, ansvarighet, transparens, och förklarbarhet). Dessa krav har på senare år tagits fram utifrån samstämmighet inom AI-samfundet kring vad algoritmer anses behöva besitta för att ingjuta förtroende, men också för att uppnå en rimlig nivå av säkerhet. *Fairness* innebär att algoritmiska beslut inte ska uppvisa några diskriminerande eller orättvisa tendenser. Kravet kräver att alla AI-system granskas noggrant för att tillse att de följer rätten till icke-diskriminering. *Accountability* innebär att det skall finnas tydliga regelverk för

vem som bär ansvaret för ett beslut som ett autonomt system fattar. *Transparency* innebär att det skall finnas tydliga svar kring målet med användning av AI i en viss kontext, och vilka delar AI:n består av. Det kan exempelvis handla om vilken data som används. Sista kravet, *explainability,* är nära besläktat med *transparency* men är mer fokuserat på att den som påverkas av ett visst beslut skall kunna förstå det algoritmiska beslutet i icke-tekniska termer (Interpol, 2020).

I maj 2017 beslutade den svenska regeringen om Sveriges digitaliseringsstrategi och definierade målet att Sverige skall vara bäst i världen på att använda digitaliseringens möjligheter (Regeringskansliet, u.å.). Strategin omfattar fem delmål som rör kompetens, trygghet, innovation, ledning och infrastruktur. I spåren av digitaliseringsstrategin publicerades också dokumentet Nationell inriktning för artificiell intelligens, med målsättningen att "Sverige skall vara ledande i att ta tillvara möjligheterna som användning av AI kan ge, med syfte att stärka både den svenska välfärden och den svenska konkurrenskraften" (Regeringskansliet, 2018). Anledningen till dokumentets framkomst motiverades av att AI är ett digitalt område som utvecklas snabbt. Kansliet menar också att för att Sverige skall vara bäst i världen på att använda digitaliseringens möjligheter så krävs en tydlig riktning för kommande prioriteringar. Det svenska regeringskansliet framhäver AI:s potential till ökad ekonomisk tillväxt och tror även att tekniken kan bidra till att lösa en rad miljömässiga- och sociala samhällsutmaningar. Dokumentet tar upp bristen på AI-kompetens i Sverige, och uppmuntrar till starkare inkorporering av AI vad gäller utbildning, forskning, innovation och användning, samt ramverk och infrastruktur, för att fullt kunna dra nytta av fördelarna AI för med sig. Därtill rekommenderas testprojekt för utveckling av AI-applikationer inom offentlig såväl som privat sektor, samt partnerskap och samarbete med andra länder kring användningen, särskilt inom EU.

Regeringen bedömer att Sverige behöver en stark grundforskning och tillämpad forskning inom AI för att säkerställa kunskaps- och kompetensförsörjningen inom området (Regeringskansliet, 2018). Rapporten är kopplad till ett forskningsprojekt i sådan anda. Projektet heter *AI: A New Scientific Revolution?* och ingår i *The Digital STS Hub* samt *The Wallenberg Autonomous Systems Program—Humanities and Society* (WASP-HS). WASP-HS stödjer forskning kring hur nya digitala verktygs framfart påverkar samhället och har ett särskilt fokus på etiska implikationer av tekniken (WASP-HS, u.å.).

Bedömningen av de risker och möjligheter som finns med AI inom brottsbekämpning är ett underbeforskat område. Trots de möjligheter som skapats med hjälp av AI omges den snabba utvecklingen av potentiella risker. Det har bland annat konstaterats att somliga grupper missgynnats vid inrättande av sådan teknik (Hong & Williams, 2019). Dessutom kan en utveckling av AI-applikationer hos stat och företag innebära en risk för inskränkning av mänskliga rättigheter om inte säkerhetsregler upprättas (OHCHR, u.å.).

Efter en initial genomsökning av databaser med forskningsartiklar ter sig den tillgängliga litteraturen kring algoritmer och AI inom brottsbekämpning tämligen rik. Däremot är den litteratur som tydligt sammanställer de risker och möjligheter som finns med dessa metoder begränsad. I och med att det är ett växande område som, i takt med

teknikens utveckling, rimligtvis kommer bli allt mer relevant är det intressant med en sammanställning av de risker och möjligheter som dagens litteratur lyfter fram. Detta kommer gynna praktiker och beslutsfattare att genomföra forskning och beslut som är mer grundade i dagens forskning. Mot bakgrund av detta valdes att genomföra en systematisk litteraturgenomgång som redogör för vad dagens forskning säger om möjligheter och risker inom brottsbekämpande arbete.

Rapporten behandlar artificiell intelligens, ett koncept som ännu saknar en entydig definition eller allmänt vedertagen avgränsning. Uppfattningen av vad som är AI har följaktligen ändrats över tid. En formulering är att AI är maskiner programmerade till att tänka som människor och imitera deras handlingar (Frankenfield, 2021). Trots att beskrivningen är talande för vad AI syftar till och kan komma att innebära i framtiden, är skildringen fortfarande en bit ifrån vad AI innebär i vetenskaplig bemärkelse idag. I vetenskaplig kontext beskrivs AI som den egenskap hos maskiner som imiterar mänsklig intelligens, karaktäriserad av beteenden såsom kognitiv förmåga, minne, lärande och beslutsfattande (Chen & Wong, 2019). Trots det existerar en diskrepans i hur AI används inom vetenskapen, där graden av komplexitet varierar. Med hänsyn till den diffusa skiljelinjen för vad som är AI kommer denna studie innefatta forskning som refererar till AI, men också algoritmer. I rapporten kommer algoritmer och AI användas alternerande, och referenser till annan litteratur nyttjar samma termer som använts i den ursprungliga forskningen.

Resterande del av rapporten har strukturerats på följande vis: nästkommande del behandlar genomförandet av metoden samt en redogörelse för den data som utvunnits. Därefter presenteras forskningens resultat, vilket följs av en diskussion kring resultatet. Detta syntetiseras och redovisas sedan i form av en slutsats. Avslutningsvis ges en reflektion kring arbetets koppling till hållbarhet och etik.

2. Metod

I följande avsnitt presenteras hur sökningen genomförts och hur filtreringen av artiklar gått till, samt en sammanställning av de artiklar som togs med.

2.1. Genomförande av systematisk litteraturgenomgång

Studien har genomförts i sju steg utifrån den struktur och tillvägagångssätt som Denscombe (2018) skriver om i *Forskningshandboken,* kombinerat med Karolinska Institutets guide för hur en systematisk litteraturöversikt görs i ett examensarbete (KI, 2022). Som första steg har studiens omfattning definierats, med syfte att klargöra vilket fält och vilka typer av studier som är av relevans för studien. Utifrån syftet definierades området som "*risker och möjligheter med algoritmer och artificiell intelligens inom brottsbekämpande arbete*".

Som andra steg specificerades villkoren för sökprocessen. Sökningarna valdes att göras i databasen Scopus, med orden "*law enforcement", "policing", "crime prevention", "crime reduction", "surveillance"* i kombination med *"algorithm"* eller *"artificial intelligence"*. Orden söktes efter i sammandrag (abstract), rubrik och nyckelord i artiklar som publicerades under tidsperioden 2018–2022. Tidsperioden valdes baserat på

forskningsområdets ålder. Anledningen till detta är att av artiklarna som kommer fram med våra sökord fr.o.m. år 2010 är fler än hälften publicerade efter år 2018, vilket tyder på att området har blivit allt mer aktuellt de senaste åren. Området utvecklas snabbt, och artiklarna inom det valda intervallet har därför hög sannolikhet att fortfarande vara relevanta. Genom att använda artiklar publicerade innan vårt valda tidsintervall ökar risken för att ta med kunskap som idag är förlegad. En av studiens medförfattare genomförde sökningen självständigt, för att säkerställa att processen genomfördes på ett konsekvent sätt. Sökningarna genomfördes under tidsperioden 8–25 april, 2022.

Som tredje steg genomfördes en kvalitetsbedömning, där artiklarna erhållna ur sökningen bedömdes utifrån deras relevans. Med motiveringen att forskningsområdet är nytt och framväxande bedömdes källorna inte utifrån antalet citeringar, eftersom nya artiklar naturligt har få citeringar. Källornas relevans bedömdes däremot i två etapper, först genom en bedömning av rubriksättningen, där de rubriker som saknade koppling till det definierade ämnesområdet sorterades bort. I andra rundan granskades artiklarnas sammandrag där de bedömdes relevanta enbart om artikeln ansågs behandla risker eller möjligheter med antingen AI eller algoritmer inom området brottsbekämpning. Först genomfördes granskningen av rubriker och artiklar av en av medförfattarna till studien. Sedan genomförde studiens medförfattare ytterligare en granskning, oberoende av den tidigare. Detta för att minimera felmarginalen vid val av artiklar.

Det fjärde och femte stegen innebar en dokumentation av studiens artiklar. Antalet artiklar som inkluderats respektive exkluderats efter relevansbedömningen dokumenterades i form av ett flödesschema. Vidare upprättades ett formulär med grundläggande information kring de valda studierna, där författare, publikationsår, journal och typ av studie dokumenterades.

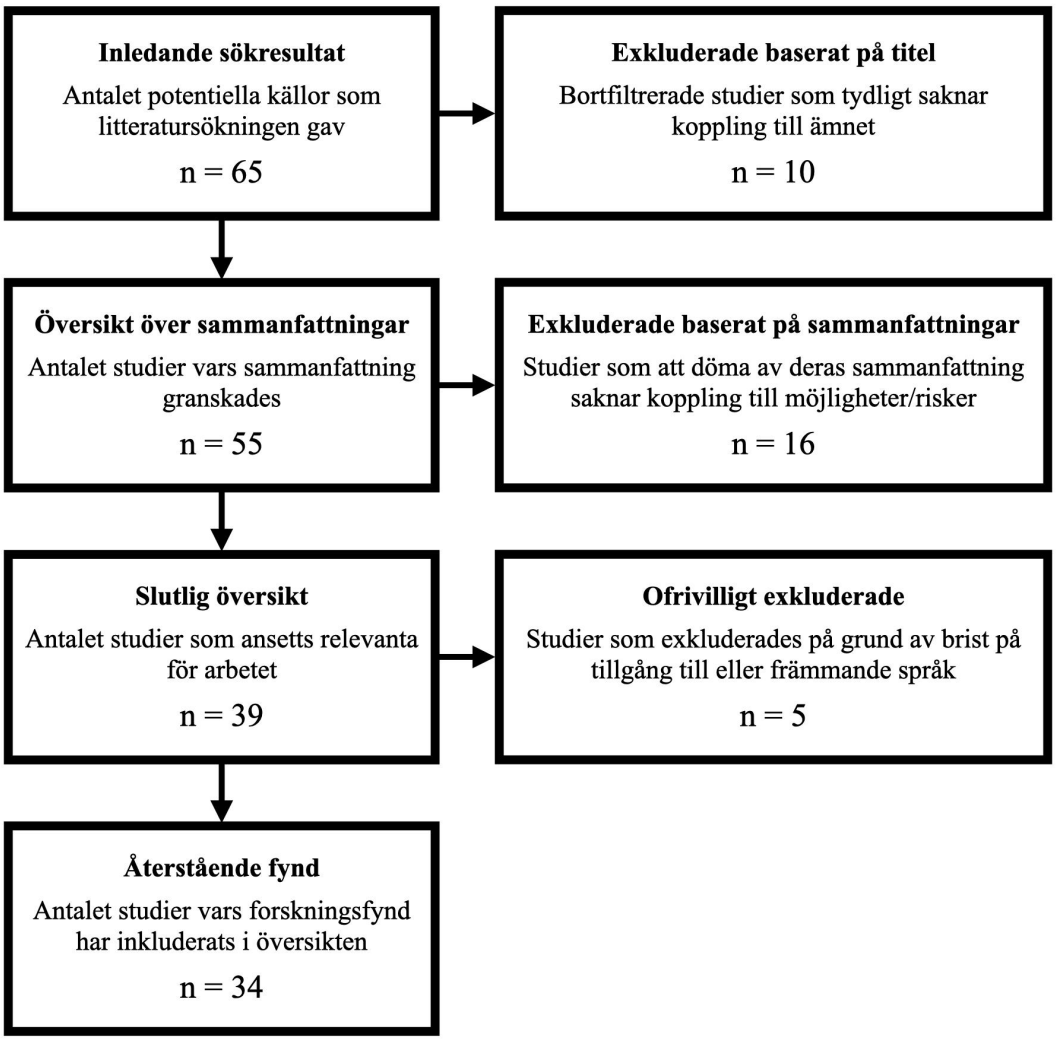
Det sjätte steget grundar sig i att genomföra kärnpunkten av forskningen: analysen. Eftersom området som forskats på grundas i användningen av ord, snarare än fynd i form av siffror, genomfördes en narrativ analys. En narrativ analys innebär att läsaren guidas genom en berättelse av de olika fynden från studierna som granskats. I jämförelse med att göra en metaanalys lämpar sig en narrativ analys när forskningsfynden inte går att kvantifiera. Narrativet valdes att förmedlas genom en tematisk analys, som innebär att forskningsfynden organiseras för att kunna identifiera genomgående teman och mönster (Mayring, 2000). Analysen som genomförts behandlar i synnerhet huruvida det råder en sammanfattad enighet eller tvetydighet kring fynden, vilket följs av en diskussion kring vad det implicerar.

Som sjunde och sista steg redovisas de slutsatser som går att utvinna ur analysen på ett sammanhängande och entydigt sätt. Slutsatsen ämnar ge en tydlig bild kring det tillstånd forskningen befinner sig i, för att på så sätt vägleda praktiker inom området var framtida studier bör ta vid.

2.2. Val av artiklar

Litteratursökningen gav ett resultat på 65 artiklar. Efter en noggrann granskning av titlar, sammanfattningar och tillgängligheten av innehåll kvarstod 34 antal artiklar som ingår i studiens resultat.

Figur 1
Litteratursökningens selektionsprocess



Kommentar. Flödesschema som visar vilka studier som av given anledning har sorterats bort ur litteraturstudien, och vilka som inkluderats.

Tabell 1
Formulär för de artiklar som ingår i studien

Författare	Publikationsår	Journal	Typ av studie
Behmer E.-J., Chandramouli K., Garrido V., Mühlenberg D., Müller D., Müller W., Pallmer D., Pérez F. J., Piatrik T., Vargas C.	2019	IFIP Advances in Information and Communication Technology	Konceptuell artikel
Bulgakova E., Bulgakov V., Trushchenkov I., Vasilev D., Kravets E.	2018	Studies in Systems, Decision and Control	Konceptuell artikel
Chang C., Chien L., Kuo E., Hwan Y.	2019	2nd IEEE International Conference on Knowledge Innovation and Invention 2019	Konceptuell artikel
Chase J., Du J., Fu N., Le T. V., Lau H. C.	2018	2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings	Empirisk studie
Contardo P., Sernani P., Falcionelli N.,	2021	CEUR Workshop Proceedings (volume	Konceptuell artikel

Dragoni A.F.		2872)	
Das P., Das A. K.	2019	Advances in Intelligent Systems and Computing	Empirisk studie
Du H., Xu Z., Yan Z., Gao S.	2018	Lecture Notes in Electrical Engineering	Konceptuell artikel
Enríquez F., Soria L. M., Álvarez-García J. A., Caparrini F. S., Velasco F., Deniz O., Vallez N.	2019	Lecture Notes in Computer Science	Konceptuell artikel
Freeman S.	2020	Proceedings of SPIE - The International Society for Optical Engineering	Empirisk studie
Ionescu B., Ghenescu M., Rastoceanu F., Roman R., Buric M.	2020	IEEE Multimedia 27(2),9116069	Konceptuell artikel
Jie Y., Liu C. Z., Li M., Choo K.-K. R., Chen L., Guo C.	2020	Applied Mathematics and Computation	Empirisk studie
Jindal S., Sharma K.	2018	Procedia Computer Science	Konceptuell artikel
Kaur B., Ahuja L., Kumar V.	2019	Proceedings - 2019 Amity International Conference on Artificial Intelligence	Empirisk studie
Kitsos P.	2020	CEUR Workshop Proceedings	Konceptuell artikel
Lanagan S., Choo K.	2021	Forensic Science International: Digital Investigation	Konceptuell artikel
Matthew U. O., Kazaure J. S., Onyebuchi A., Daniel O. O., Muhammed I. H., Okafor N. U.	2021	Proceedings of the 2020 IEEE 2nd International Conference on Cyberspace	Konceptuell artikel
Milivojevic S.	2021	Crime and punishment in the future internet: Digital frontier technologies and criminology in the twenty-first century	Konceptuell bok
Molina-Molina J. C., Salhaoui M., Guerrero-González A., Arioua M.	2021	Intelligent Sensing Systems for Vehicle	Empirisk studie
Pawlicka A., Choraś M., Przybyszewski M., Belmon L., Kozik R., Demestichas K.	2021	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	Konceptuell artikel
Rajamäki J., Sarlio-Siintola S., Simola J.	2018	European Conference on Information Warfare and Security, ECCWS	Konceptuell artikel
Rajapakshe C., Balasooriya S., Dayarathna H., Ranaweera N., Walgampaya N., Pemadasa N.	2019	2019 International Conference on Advancements in Computing, ICAC 2019	Empirisk studie
Ramirez J., Campo-Archbold A., Zapata A., Diaz-Lopez D., Pastor-Galindo J., Gomez Marmol F., Aponte J.	2022	IEEE Internet Computing	Empirisk studie
Ridzuan Khairuddin A., Alwee R., Haron H.	2020	IOP Conference Series: Materials Science and Engineering	Empirisk studie
Sherer J. A., Sterling N. L., Burger L., Banaschik M., Taal A.	2018	Advanced Sciences and Technologies for Security Applications	Konceptuell artikel
Simpson T.	2021	AIAA/IEEE Digital Avionics Systems Conference - Proceedings	Konceptuell artikel
Singh A., Anand T., Sharma S., Singh P.	2021	Proceedings of the 6th International Conference on Communication and Electronics Systems	Empirisk studie
Smith M., Miller S.	2022	AI and Society	Konceptuell artikel
Sun S.	2020	Journal Of Visual Communication and Image Representation	Empirisk studie
Suralkar S., Gangurde S., Chintakindi S., Chawla H.	2020	Lecture Notes on Data Engineering and Communications Technologies	Konceptuell artikel
Thoa Mac T., Copot C., Lin C.-Y., Hong Hai H., Ionescu C. M.	2020	Journal of Physics: Conference Series	Konceptuell artikel
Toppireddy H. K. R., Saini B., Mahajan G.	2018	Procedia Computer Science	Empirisk studie

Wang H., Ma S.	2022	Socio-Economic Planning Sciences	Empirisk studie
Wyatt A.	2021	Defence Studies	Konceptuell artikel
Zhang R.	2021	Proceedings - 2021 International Conference on Computer Information Science and Artificial Intelligence, CISAI 2021	Konceptuell artikel

Kommentar. De artiklar som inkluderats i studien, sorterade enligt författarnas bokstavsordning.

Utifrån de återstående fynden från granskningen av de insamlade källorna identifierades gemensamma teman. Dessa teman presenteras i nästkommande resultatavsnitt.

3. Resultat

Arbetet resulterade i ett antal möjligheter och risker associerade med användandet av AI inom brottsbekämpning. För att på bästa sätt återge möjligheterna på ett begripligt sätt för läsaren så har de delats in i tre olika områden. Vissa algoritmer används för att förutse och förhindra att brott begås överhuvudtaget, och dessa redovisas i avsnittet *Före brott*. Andra är gjorda för att assistera vid och upptäcka pågående brott, och dessa redogörs för i avsnittet *Under brott*. Sist kan algoritmer också tjäna vid brottsutredning väl efter brottet har begåtts, och dessa faller under rubriken *Efter brott*. Slutligen presenteras riskerna som tar fokus på de etiska aspekter som presenterats i de studerade källorna.

Figur 2

Sammanställning av identifierade möjligheter och risker

Möjligheter	Före brott	Övervakning av högriskindivider Identifiering av objekt och rörelser Språkanalys för beteendeprognostisering Förutsägelser och prognoser
	Under brott	Gruppbeteendeanalys och långvarig övervakning från ovan Trafikinformation i realtid och identifiering av registreringsskyltar Permanent övervakning av marina skyddsområden Situationsanalys i realtid
	Efter brott	Drönare som förlängd arm Intelligent identifiering Lögndetektering vid förhör Genomsökning av hårddiskar Mönster och kartläggning
Risker	Diskriminerande tendenser Brist på personligt ansvar Massövervakning Urholkning av privatliv och integritet	

3.1. Möjligheter

I detta avsnitt behandlas de möjligheter som de olika artiklarna tar upp. Möjligheterna presenteras i tre olika underdelar baserat på skeendet kring ett brott, nämligen perioden före brott, perioden under brott och perioden efter brott.

3.1.1. Före brott

I artiklarna har fyra olika typer av brottsförutsägande och brottsförebyggande tillvägagångssätt identifierats: *Övervakning av högriskindivider*, *Identifiering av objekt och rörelser*, *Språkanalys för beteendeprognostisering*, samt *Brottsförutsägelser och prognoser*.

3.1.1.1. Övervakning av högriskindivider

Flera artiklar ser stor potential i att använda algoritmer för att identifiera och övervaka personer av intresse. Fysiska markörer som ansikten och fingeravtryck kan kännas igen och delas mellan avdelningar och regioner för att bevaka personer som bedöms vara brottsbenägna (Contardo et al., 2021; Pawlicka et al., 2021; Smith & Miller, 2022; Wyatt, 2021).

Ett potentiellt mer kraftfullt spår är digitala fotavtryck. De ger brottsbekämpningen möjlighet att identifiera potentiella brottslingar som saknar tidigare misstanke. Exempelvis kan sådana verktyg identifiera individer som uppviglar till brott vid planerade demonstrationer (Ramirez et al., 2022) men de kan också upptäcka personer med psykologiska åkommor farliga för dem själva och andra (Bulgakova et al., 2019; Ionescu et al., 2020; Jindal & Sharma, 2018).

3.1.1.2. Identifiering av objekt och rörelser

En annan domän där datorers bildförståelse kommer till användning är övervakningsfilm. Med hjälp av *computer vision* (sv. datorseende) kan enorma mängder videomaterial kontrolleras för misstänksamma aktörer (Ionescu et al., 2020; Smith & Miller, 2022), föremål (Enríquez et al., 2019; Ionescu et al., 2020), och ageranden (Rajapakshe et al., 2019). Det finns exempelvis verktyg för att urskönja och följa individer i rörelse, men också verktyg som studerar gruppers ageranden (Ionescu et al., 2020; Rajapakshe et al., 2019). Det utvecklas även AI-program som gör det möjligt att upptäcka vapen (Enríquez et al., 2019; Singh et al., 2021). Oavsett vad de är gjorda för att upptäcka besitter de vissa inherenta egenskaper som eftersöks hos övervakare. De är snabba, outtröttliga, och diskreta (Wyatt, 2021).

3.1.1.3. Språkanalys för beteendeprognostisering

Flera studier har visat att det är möjligt att med hjälp av språkanalys förutse både planerade och oplanerade dåd innan de sker (Ionescu et al., 2020; Jindal & Sharma, 2018; Ramirez et al., 2022). En populär datakälla är sociala medier. Där finns offentlig kommunikation och i vissa fall även privata samtal att tillgå. Algoritmer gör det möjligt att analysera de enorma datamängderna (Behmer et al., 2019). Inläggen matas in till algoritmer, som i grova drag kan avgöra avsändarens sentiment. Detta görs bland annat genom att algoritmen identifierar information som består av kombinationer av namn, platser, organisationer, och tid. Användares beteenden på plattformen kan aggregeras och jämföras mot tidigare brottslingars aktiviteter. På så vis kan digitala profiler

konstrueras vilka kan studeras mer ingående av mänskliga utredare. Det kan röra sig om personer som lider av svåra psykiska åkommor (Jindal & Sharma, 2018), men profilerna kan också användas för att upptäcka grupper som planerar illdåd (Ramirez et al., 2022).

Språkanalys behöver dock inte begränsas till sociala medier. Nyttan med avlyssning blir mycket större tack vare nya mjukvaruredskap. Det är nu möjligt att koppla röster till personer, konvertera tal till text, och urskönja ord från dåliga inspelningar (Ionescu et al., 2020). Skulle mikrofoner inte finnas tillgängliga är det till och med möjligt för datorer att läsa läppar från videor (Ionescu et al., 2020).

3.1.1.4. Brottsförutsägelser och prognoser

De studier som studerats i arbetet ger flera exempel på metoder att använda för att förutsäga under vilka omständigheter nya brott sker. Det nämns hur högriskindivider identifieras genom att sammanställa beteendemässiga data från olika datakällor och analysera denna datasamling (Freeman, 2020; Kaur et al., 2019; Smith & Miller, 2022; Zhang, 2021). Med hjälp av algoritmer som tränats genom *unsupervised learning* (sv. oövervakad maskininlärning) går det att notera avvikelser från individers beteendemönster, varpå dessa kan klassificeras i olika grupper baserat på hur riskabelt det avvikande beteendet, enligt algoritmen, bedöms vara (Jindal & Sharma, 2018). Det nämns även hur AI tillför en ny dimension av träffsäkerhet vid prediktion; tekniken bygger på icke-linjära funktioner vilket gör det möjligt att upptäcka icke-linjära mönster (Ridzuan Khairuddin et al., 2020). Däremot krävs förbättringar inom beslutsfattande, bearbetning och hantering av *big data* (sv. stordata), samt en specifik IT-infrastruktur för att göra analyser när datamängden blir stora (Bulgakova et al., 2019). Om infrastrukturen i framtiden skulle komma att falla på plats beräknas polismyndigheterna få ett ökat ansvar för hanteringen av dessa, ofta privata, datamängder (Sherer et al., 2018).

Idag går det att analysera onormalt beteende hos individer och grupper för att prognostisera om det finns risker för brott i ett område genom informationsinsamling från konversationer (Ionescu et al., 2020). Det är dock först när datan analyserats som det går att dra adekvata och tillräckligt väldefinierade slusatser för att förutsäga var nästa brott kommer ske (Toppireddy et al., 2018). I sådana processer blir en automatiserad videoövervakningsprocess en av grundpelarna för att kunna förutsäga framtida brott (Rajapakshe et al., 2019). En underlättande aspekt i den frågan baseras då på tillgången till data om tidigare brott (Das & Das, 2019). Med tillräckliga brottsdata kan sedan AI förbättra polisens resursallokering och optimera för faktorer som responstider (Chase et al., 2018) eller arresteringar (Jie et al., 2020).

3.1.2. Under brott

Ur artiklarna har följande användningsområden identifierats, kopplat till nyttjande av AI och algoritmer under tiden då ett brott begås: *Gruppbeteendeanalys och långvarig övervakning från ovan*, *Trafikinformation i realtid och identifiering av registreringsskyltar*, *Permanent övervakning av marina skyddsområden,* samt *Situationsanalys i realtid*.

3.1.2.1. Gruppbeteendeanalys och långvarig övervakning från ovan

Ett användningsområde som identifierats av Simpson (2021) är övervakning av folkmassor i realtid. Med hjälp av AI kan polis analysera massbeteenden under såväl fredliga som våldsamma sammankomster, där AI automatiskt identifierar faktorer som gruppdensitet, hastigheter, förflyttningsbeteende, samt identifiera de primära våldsverkarna. En fördel som lyfts fram med detta är att det sparar resurser, då polis kan placera ut sin personal på mer lämpliga platser till följd av det bättre informationsunderlaget som drönaren i detta fall skapat.

Användning av autonoma system, som drönare, möjliggör enligt Wyatt (2021) även för bevakning under en längre tid då systemet till stor del agerar på egen hand med begränsad medverkan från en människa. Alternativet är att placera ut spaningsgrupper men dessa kan inte verka under lika lång tid, utsätts för personlig risk, och är svårare att hålla dolda från kontraspionage. Drönare med inbyggd analys av inspelat material med hjälp av AI är alltså något som utforskas, men även mindre sofistikerade drönare som endast inhämtar material utan egen analys är användbara i och med framtåget av mer komplexa metoder för videoanalys med hjälp av AI. Det autonoma systemet kan därför användas enbart för informationsinhämtning, för att sedan överlämna materialet till andra system som i sin tur genomför analyser på den stora datamängd som inhämtats för att hitta nyttig information ifrån den (Wyatt, 2021).

Även mänskliga fel vid granskning av videomaterial kan elimineras när övervakning bedrivs av drönare med AI, enligt Thoa Mac et al. (2020). Genomgång av stora mängder videomaterial för att finna vad som behövs för exempelvis en utredning tar lång tid och det sker ofta misstag på grund av den enformiga och långdragna arbetsprocessen. Det kostar även stora resurser i form av arbetskraft, vilket skulle kunna reduceras kraftigt med hjälp av AI som filtrerar videomaterialet i realtid, utan några misstag.

3.1.2.2. Trafikinformation i realtid och identifiering av registreringsskyltar

Ett annat tillämpningsområde av drönare som lyfts fram är övervakning av trafik (Chang et al., 2019; Du et al. 2018; Thoa Mac et al., 2020). Med allt fler bilar i våra städer uppstår problem kring hur man bäst handskas med trängsel och trafiksäkerhet. Idag används bland annat vanliga övervakningskameror och infraröd teknik för att inhämta information kring fordons hastighet och trafikflödet. Problemet är att det är svårt att avgöra var dessa verktyg skall placeras ut, samt att de är oflexibla när de väl placerats. Därav kan inte alltid den nödvändiga trafikinformationen inhämtas. En lösning som Thoa Mac et al. (2020) lyfter är att nyttja drönare som med hjälp av bildsensorer kan överföra trafikinformation i realtid till en kontrollstation. Övervakningen blir mer flexibel och kan täcka ett större övervakningsområde än vad enskilda utplacerade kameror kan. En konsekvens av sådana system är till exempel att information om att ett fordon överskrider hastighetsbegränsningen automatiskt inkommer till polis som sedan kan agera (Thoa Mac et al., 2020).

Flera artiklar belyser möjligheterna som finns med att utplacerade övervakningskameror kan identifiera registreringsskyltar på fordon (Chang et al., 2019; Du et al. 2018). En möjlighet som utforskas är att polis skall kunna ge ett övervakningssystem ett registreringsnummer som input och därefter låta systemet spåra hur fordonet rör sig i

realtid genom att använda video från alla kameror som ingår i systemet. Genom att låta systemet förutspå vart fordonet är på väg kan polis förbereda en arrestering på lämpligt område (Chang et al., 2019; Du et al. 2018).

3.1.2.3. Permanent övervakning av marina skyddsområden

Ett annat tillämpningsområde av autonoma system är inom den marina miljön (Molina-Molina et al., 2021; Rajamäki et al., 2018). Molina-Molina et al. (2021) föreslår bland annat hur robotar kan användas i syfte att bedriva permanent övervakning inom marina skyddsområden genom att med hjälp av AI identifiera fartyg som bedriver misstänksam verksamhet. Det lyfts dock att det i marina miljöer är svårt att försäkra att mätvärdena är korrekta och att resultaten därför inte ska betraktas som fullt pålitliga. Som komplement till dessa typer av system och andra övervakningsverktyg, som till exempel helikoptrar, kan även drönare komma till användning vid informationsinhämtning (Rajamäki et al., 2018).

3.1.2.4. Situationsanalys i realtid

Något som utforskas är hur intelligenta system kan analysera situationer i realtid (Enríquez et al., 2019; Suralkar et al., 2020). Enríquez et al. (2019) utforskar hur realtidsrådgivning till potentiella offer kan ske när övervakningskameror med hjälp av AI uppfattar eventuella hot på video som exempelvis vapen. Användare rekommenderas ha en applikation installerad på sin telefon som rådger användaren kring hur denne bör agera, beroende på dess position i relation till det potentiella hotet och andra parametrar som till exempel var panikartad flykt uppstått. Dessa råd kan handla om vilka flyktvägar som är lämpligast eller om personen bör låsa in sig på plats. Systemet möjliggörs genom att integrera övervakningsmaterial med ett subsystem som låter användarnas telefoner dela information av gemensamt intresse med varandra, också känt som *crowdsensing*. Enligt författarna av artikeln är utmaningen att utveckla en metod som kan reagera snabbt och korrekt i ett ständigt föränderligt scenario samtidigt som flera parametrar tas i beaktande.

Suralkar et al. (2020) utforskar hur drönare med integrerad AI skall kunna sätta en händelse i en kontext. Det skulle innebära att drönaren själv, i realtid, kan avgöra om det som observeras faktiskt är ett pågående brott, eller ett falskt positivt resultat. Detta till skillnad från många andra övervakningssystem med AI som kan avgöra om det är en våldsam handling som tar plats, men inte om det exempelvis rör sig om en egentligen oskyldig handling på grund av kontexten (Suralkar et al., 2020).

3.1.3. Efter brott

Ur litteraturen har fyra områden identifierats där AI och algoritmer skapar möjligheter efter brottet har ägt rum. De är: *Drönare som förlängd arm*, *Intelligent identifiering av individer och föremål*, *Lögndetektering vid förhör*, *Genomsökning av hårddiskar*, samt *Mönster och kartläggning*.

3.1.3.1. Drönare som förlängd arm

Drönare har enligt litteraturen en stor potential för att, i kombination med artificiell intelligens och maskininlärning, tjäna polisiära syften (Contardo et al., 2021; Matthew et al., 2021; Simpson, 2021). Rättsväsendet har under senare tid upplevt ett växande problem när det kommer till att hantera den dynamiska natur som publika

sammankomster ofta karaktäriseras av. Utöver att drönare i kombination med artificiell intelligens gör det möjligt att i realtid identifiera brottslig aktivitet, så skapar appliceringen också värdefullt bevismaterial vid brottsutredning. Genom inkorporeringen av AI kan de självständigt interagera med luft- och marknavigationssystem, vilket spelar en stor roll för möjligheten till bevisinsamling även under svåra förhållanden (Simpson, 2021). Den autonoma egenskapen innebär att drönarna kan verka bortom siktlinjer; en aspekt som uppges vara en stor fördel då systemet tillåts tjänstgöra utan direkt kontroll av människan (Matthew et al., 2021).

Förutom att dokumentera publika sammankomster kan inkorporering av AI i videoupptagningen innebära möjligheten att identifiera och lokalisera brottslingar, samt assistera vid räddnings- och sökoperationer (Matthew et al., 2021). Detta förverkligas genom att dra nytta av Internet of Things (IoT) (sv. sakernas internet) samt använda sig av en specifik gren av artificiell intelligens som är vanlig för bildanalys: *convolutional neural networks* (CNNs) (sv. konvolutionella neurala nätverk) (Contardo et al., 2021). Videoanalysen inkorporerad i drönare antas ha ett stort värde för länder som lider av en hög närvaro av kriminella nätverk (Matthew et al., 2021). Användningen av drönare och appliceringen av AI kan dessutom gagna utvärdering av polisens ingripande vid folksamlingar (Simpson, 2021).

3.1.3.2. Intelligent identifiering av individer och föremål

Övervakningskameror har sedan kamerans uppkomst varit ett viktigt polisiärt verktyg för att kunna identifiera individer som rörts sig i ett visst område kring en kritisk tidpunkt. En av de vanligaste biometriska teknikerna är ansiktsigenkänning, vilket har varit ett attraktivt forskningsområde i så länge som 40 år (Contardo et al., 2021). Ansiktsigenkänning bär fördelen att det kan användas passivt, vilket innebär att personen som önskas identifieras inte behöver vara medveten om användandet. Tidigare tekniker har varit bristande i deras effektivitet på bild och video från övervakningskameror, där förutsättningarna för bilden inte är ideala. AI och maskininlärning förbättrar igenkänningstekniken, där bildanalysteknik baserad på CNN är bra på att känna igen ansikten även under svåra förhållanden (Contardo et al., 2021). Dessutom utnyttjas algoritmer för att återställa och förbättra bildkvalitéten vid oskarpa bilder hämtade från videoupptagningar (Sun, 2020). Genom analys av kameror och nätverk av dessa kan tekniken analysera närvaro av personer, vilket skapar en möjlighet att vid sökning av en person få en redogörelse av individens historiska händelsemönster och beteende (Ionescu et al., 2020).

En annan biometrisk uppgift som sedan en lång tid tillbaka varit föremål för analys är fingeravtryck. Litteraturen förklarar att traditionella algoritmer har fungerat väl på avtryck som är insamlade med avsikt, men presterat sämre på latenta fingeravtryck, alltså sådana som oavsiktligt lämnats på ytor (Contardo et al., 2021). Lösningen på problemet förväntas ligga i att utveckla tekniken med hjälp av CNN:s.

Andra möjligheter som AI ger är videoanalys av registreringsskyltar. Genom att designa intelligenta system som kopplar samman data från olika övervakningskameror skall en sökning på ett visst registreringsnummer ge polisen en notis när registreringsskylten upptäcks från någon av kamerorna (Chang et al., 2019). Polis kan då få en god

uppfattning om vart den misstänkte är på väg och förbereda en arrestering på lämplig plats.

3.1.3.3. Lögndetektering vid förhör

I brottsutredningar är verbala uppgifter en essentiell del av processen för att lösa brott. En svårighet med denna typ av bevismaterial, vare sig det gäller vittnesmål eller förhör med den brottsmisstänkte, är att avgöra sanningshalten i de uttalanden som görs. Litteraturen visar hur AI-baserad analysering av tal kan bedöma huruvida personen som förhörs ljuger eller talar sanning (Ionescu et al., 2020; Pawlicka et al., 2021). Ett särskilt exempel från litteraturen bygger på två moduler som kopplas samman för att avgöra trovärdigheten (Ionescu et al., 2020). Den första delen är en analys av emotionella tillstånd, där data från bild och ljudupptagningar analyseras med hjälp av bland annat *deep neural networks* (sv. djupa neurala nätverk) och avgör intervjuobjektets sinnestillstånd. Deep neural network är ett slags nätverk av noder vars struktur inspirerats av människohjärnan som i vissa situationer möjliggör en mycket effektiv och snabb maskininlärning. Den andra delen utgör en analys av psykologiskt tillstånd och innebär realtidsanalys av hjärtslag och andningsfrekvens. De kroppsliga mätvärdena utvinns endast ur videoanalys, vilket innebär att lögndetekteringen kan göras utan att intervjuobjektet är medveten om det.

3.1.3.4. Genomsökning av hårddiskar

Lanagan & Choo (2021) diskuterar möjligheten för genomsökning av hårddiskar för att hitta olagligt material såsom barnpornografi. Det huvudsakliga användningsområdet som identifieras är vid gränskontroller. I artikeln föreslås det att man kan söka efter specifika indikationer på barnpornografi, som att hårddisken till stor del innehåller videofiler. Detta kan i sin tur ligga till grund för en mer omfattande genomsökning. Fördelen med användningen av AI anses alltså vara att hårddisken kan genomsökas på ett effektivt sätt, samt att dess ägares personliga integritet inte kränks lika mycket som ifall en människa utför sökningen.

3.1.3.5. Mönster och kartläggning

För att få ut betydelse av big data krävs inkorporering av artificiell intelligens (Pawlicka et al., 2021). Ett exempel som lyfts är hur AI kan klara upp seriebrottslighet genom att upptäcka icke triviala samband. Med hjälp av AI förväntas det även bli enklare att nå kriminella nätverk, då möjligheten att identifiera individer kopplade till kriminella personer blir enklare (Sherer et al., 2018). Det påpekas däremot att associerade individer inte är intressanta i sig, utan de individer som bör studeras är de som har koppling till den kriminella aktiviteten. Bulgakova et al. (2019) betonar vikten av infrastruktur för att kunna använda sig av big data i utredningar. Vidare lyfter Sherer et al. (2018) hur domare och nämndemän kommer se på trovärdigheten i de fynd som kan presenteras med hjälp av big data som en risk. Med hjälp av AI-analytiker kan mänskliga analytikers arbete effektiviseras, vilket tillåter människan att lägga större fokus på effektiv respons snarare än utredning (Wyatt, 2021). AI bär också potentialen att reda ut mer alldagliga brott som bedrägerier och småstölder mer effektivt än idag (Pawlicka et al., 2021).

3.2. Risker

Ur litteraturen gick det att urskilja fyra områden som beträffar risker associerade till AI och algoritmer inom brottsbekämpning. De risker som identifierades var *Diskriminerande tendenser, Brist på personligt ansvar, Massövervakning,* samt *Urholkning av privatliv och integritet.*

3.2.1. Diskriminerande tendenser

Litteraturen framhäver hur användningen av algoritmer i polisarbete riskerar att upprätthålla diskriminering av olika folkgrupper, i synnerhet baserat på etnicitet (Kitsos, 2020; Milivojevic, 2021). Ett exempel på detta som Kitsos (2020) tar upp är vid prediktivt polisarbete, där algoritmer ämnade att samordna polisresurser tränas upp med hjälp av tidigare data från poliser, som alltså redan kan ha vissa diskriminerande tendenser. I artikeln så citeras en studie där det visade sig att unga svarta män löper större risk att bli stoppade av polisen än unga vita män då algoritmer används för samordning av polisstyrkor, just på grund av den historiska datan som används. En metod som Zhang (2021) föreslår för att motverka detta är att låta en tredje part väga in på hur data som samlas in från algoritmer skall tolkas och användas. Wang & Ma (2022) hävdar även att AI kommer öka social diskriminering baserad på etnicitet och kön, vilket i sin tur kan leda till ökad kriminalitet i samhället.

3.2.2. Brist på personligt ansvar

Litteraturen diskuterar även var ansvaret för algoritmers användning ligger. Zhang (2021) påpekar att det finns en risk för att oskyldiga personer pekas ut som misstänkta för brott endast för att de råkar passa in på en viss profil. I det fallet så utgör alltså en algoritm i sig grunden för en misstanke, utan att polisen själv är helt införstådd på resonemanget bakom slutsatsen som nåtts. För att förhindra detta så uppmanas poliser ha ett kritiskt förhållningssätt gentemot data och rekommendationer som erhålls från AI-program. Lanagan & Choo (2021) argumenterar även för att man i brottsbekämpningen skall använda sig av *explainable AI* (XAI) (sv. förklarbar AI), eftersom en människa annars inte kan ta juridiskt ansvar för vad algoritmen gör. Kitsos (2020) menar att polisen måste skapa en arbetskultur där de själva står ansvariga för sina handlingar. Lanagan & Choo (2021) citerar Pasquale (2019), som menar att automation i juridiska frågor riskerar att undergräva grundläggande principer om ansvar ifall det inte finns människor som ansvarar för hur de fungerar.

3.2.3. Massövervakning

Spaning och övervakning har alltid varit ett viktigt verktyg för myndigheter. Efter Snowdens avslöjanden om att amerikanska och europeiska rättsmyndigheter bedriver massövervakning över dess invånares elektroniska kommunikation har en diskussion påbörjats kring behovet av denna typ av övervakning samt hur den går till. Kombinationen av internet och digitala teknologier som använder AI har gett upphov till övervakningssystem som ständigt ser över folks liv och som resulterar i en förlust av anonymitet som aldrig tidigare skådats (Kitsos, 2020; Milivojevic, 2021). Kinas pågående projekt kring social kreditvärdighet möjliggörs av omfattande övervakning på befolkningen och syftar till att ge en individ eller företag en viss poäng baserat på hur denne uppför sig i samhället (Kitsos, 2020). Poängen avgör om du får vissa privilegier

eller ej. Detta system är ett exempel på vad som kan möjliggöras med hjälp av massövervakning.

3.2.4. Urholkning av privatliv och integritet

När polisstyrkor använder sig av AI för att kunna förutse och förebygga kriminalitet krävs högkvalitativa och tillräckliga data. Smith och Miller (2022) konstaterar att det redan finns massiva mängder användbara data insamlad och att denna informationsarsenal endast kommer att växa. De menar att problemen uppstår när datan skall tillgås. Artikelförfattarna vill kunna aggregera en rad datakällor från olika domäner. Det rör sig om öppet tillgängliga data från brottsregister (Wang & Ma, 2022) och sociala medier (Freeman, 2020), men även data från privata chattkonversationer (Bulgakova et al., 2019), affärer (Freeman, 2020; Wyatt, 2021), medicinska journaler (Freeman, 2020), och skatteregister (Smith & Miller, 2022). Heterogena datasamlingar beskrivs också som problematiska eftersom de kräver hög sofistikation, alternativt mycket manuell handpåläggning, för att sammanfoga. Ett par artiklar rekommenderar därför nationella- och internationella initiativ för att skapa breda, gemensamma standarder för datasystemen (Freeman, 2020; Pawlicka et al., 2021).

En fråga som blir relevant i relation till dessa typer av datamängder är frågan om integritet. Ofta innebär brottsbekämpning med hjälp av AI hämtning av data från flera olika källor, där det samlade underlaget målar en detaljerad bild av en individ. Den typ av data som används för brottsprognoser består i många fall av känslig information som etnicitet, hudfärg, politisk åsikt, religionstillhörighet eller annan trosuppfattning, hälsotillstånd, sexuell läggning eller biometrisk data för att identifiera en person (Kitsos, 2020). Även mindre personliga användardata såsom räkningar som betalas och digitala tjänster som används kan ingå (Zhang, 2021). Datan som godkänts för användande på en särskild plattform kan, ihop med annan data och tagen ur sitt sammanhang, måla upp en detaljerad bild av en individ, utan att individen är medveten om det. En risk med expansionen av AI kan således innebära en urholkning av privatliv och ett hot mot den personliga integriteten.

Ur litteraturen framgår det att det här sättet att använda komplement av data för att identifiera individer existerar redan idag. Det lyfts fram hur många västerländska länder har strikt reglering av telefonavlyssning, medan upptagning ur sociala medier är mindre reglerat (Rajamäki et al., 2018). Eftersom sociala medier är publika och fria att tillgå så uppfattar polis och säkerhetstjänstemän genomsökningen som berättigad (Rajamäki et al., 2018). Den stora skillnaden mellan att vem som helst och en tjänsteman tillgår datan är konsekvenserna det innebär, då genomsökning för brottsbekämpning kan få reella konsekvenser på någons privatliv och framtid.

Ett fall som lyfts fram är när företaget Clearview AI, som genom en biometrisk ansiktsigenkänningsalgoritm använde sig av bilder från sociala medier för att identifiera gärningsmän (Smith & Miller, 2022). I stället för den smala mängden bilder som finns att tillgå i nationella databaser, som passfoton och körkort, så lät de sin produkt utnyttja de miljarder bilder som finns tillgängliga på sociala medier och andra digitala plattformar. Det här mötte stark kritik från allmänheten, juridiska repressalier, och en motreaktion från företagen som tillhandahåller fotografierna i form av *cease-and-desist*-brev (Smith & Miller, 2022). De menade på att Clearview AI genom

sina tjänster hade brutit mot användaravtalet. Det konstateras att den snabba utvecklingen av datatillgänglighet i kombination med de ökade användningsområdena skapar ett behov av nya strukturer för medgivande och datasäkerhet (Smith & Miller, 2022). Vidare anmärker Rajamäki et al. (2018) att akademiska och offentliga debatter har kommit att skifta fokus från datainsamling till dataanalys och dataanvändning. Ett skifte som, med föregående exempel i åtanke, kommer påverka möjligheten att skydda individens privatliv.

4. Diskussion

Två sorters artiklar förekommer mest frekvent i litteraturen. Vissa förespråkar implementering av digital teknik, och andra föreslår ramverk för digital teknik. Artiklarna är i huvudsak vänligt inställda till den behandlade tekniken och fokuserar främst på de positiva aspekterna med en implementering av AI, snarare än de negativa. De artiklar som berör samma ämnen är tillika samstämmiga i frågor om dagens förmågor, teknikens potential, och vad framtida arbeten bör behandla.

Eftersom merparten av artiklarna uttrycker praktisk nytta som ett mål är det anmärkningsvärt hur få av de presenterade lösningarna som faktiskt prövats i fält. Inte heller i arbetenas bakgrund framgår det huruvida det finns en konkret efterfrågan på deras lösningar. Detta skulle kunna tyda på att det råder en dissonans mellan de digitala verktyg som brottsbekämpningen söker använda och den publika forskningen.

Ett genomgående tema för artiklarna är förhoppningen för vad en alltmer omfattande implementering av AI inom brottsbekämpning skulle kunna åstadkomma. När det kommer till tekniska framsteg så är det en optimistisk bild som målas upp, med allt från drönare som kan analysera videomaterial i realtid till scanning av sociala medier för att upptäcka brott och stoppa dem innan de sker.

Ett sätt att tolka litteraturen kan vara att framtiden för brottsbekämpning kommer präglas av extremt högteknologisk AI, och det råder ingen tvekan om att tekniken kommer att förbättras. Emellertid innebär den begränsade informationen kring faktisk implementering eller realiserbarhet av de testmodeller och ramverk som en del artiklar föreslår att det blir svårt att avgöra om de bör ses som möjligheter, eller om forskningen är i ett för tidigt skede för att kunna göra den bedömningen. Samtidigt är det viktigt att forskning som befinner sig i ett tidigt skede publiceras, då den kan belysa nya forskningsområden.

De användningsområden som i resultatet identifierats i avsnittet *Före brott* kretsar främst kring bild- och textanalys. En observation är att bild- och textanalysen i första hand ämnar identifiera beteenden, som i sin tur indikerar att ett brott skall begås. De studier som ingår i studien bygger således på förutsättningen att det går att dra kopplingar mellan ett visst beteende, som inte är ett brott, och att ett faktiskt brott skall ske. Studierna presenterar inte förslag på hur sådana typer av prediktioner skall användas, alltså hur en individ som bedöms ha ett riskbeteende skall bemötas. Det hade varit befogat med en diskussion kring hur den sortens underlag bör användas av polis, då det faktiskt handlar om att korrigera och peka ut människor som ännu inte har begått

ett brott. Följaktligen går det även att fråga sig huruvida en sådan implementering i själva verket skulle begränsa ramarna för vad som anses tillåtet i samhället, då en större mängd handlingar (icke-brottsliga) ändå skulle kunna väcka polisens intresse och ligga till grund för misstanke.

Den här kategorin präglas dessutom av problemet att det inte finns något facit som prognosen kan jämföras mot. Teknikens syfte är att förutsäga brott, för att på så vis kunna förhindra att de begås. Det innebär att det inte kommer finnas några verkliga data att jämföra prediktionerna med. Det blir svårt att bekräfta att det beteende som har identifierats, och sedan korrigerats, faktiskt skulle leda till ett brott om det inte hade upptäckts, eller om det i själva verket var ett falskt-positivt utslag. Det går av den anledningen att anta att resursen bör kunna nyttjas av polisen i samma mån som andra obekräftade tips.

Litteratur utanför den genomförda forskningen lyfter fram *automation complacency* (sv. automationspassivitet) som ett fenomen där personer som använder sig av algoritmbaserade hjälpmedel vid beslutsfattande blir passiva och inte ifrågasätter tekniken (Adensamer et al., 2021). I samma studie görs en distinktion mellan termerna *automation complacency* och *automation bias* (sv. automationsbias), som istället beskriver en tendens att lita på förslagen som dessa hjälpmedel ger. En potentiell risk med det blir således att människan vid inrättande av algoritm- eller AI-baserad teknik gömmer sig bakom algoritmerna och på så sätt avskriver sig ansvaret från konsekvenserna. Litteraturen i vår genomförda studie verkar däremot sakna en diskussion om huruvida detta fenomen kan komma att påverka de som bär ansvaret för användningen av AI inom brottsbekämpning.

De användningsområden som presenterats i avsnittet *Under brott* utgick huvudsakligen i att genomföra någon typ av bildanalys, antingen med videoupptagningar från stationära kameror, eller från autonoma system (drönare), som alltså tillför en ökad bredd för var tekniken kan komma att nyttjas. Avsikten med bildanalys är främst att kunna upptäcka brott i realtid för att i första hand kunna stoppa dem, men även att assistera potentiella brottsoffer med till exempel flyktvägar. Även då litteraturen listar hoppfulla förslag på hur bildanalys kan komma att gynna brottsbekämpning finns en problematik med hur inrättandet av sådan teknik kan komma att bidra till en form av massövervakning.

Ett övervakningsområde utan lika uppenbar problematik med hänsyn till medborgarövervakning är övervakning av marina skyddsområden. Inrättande av den typ av teknik till havs skulle potentiellt uppfattas som mindre riskfyllt, då det främst innebär övervakning av områden där invånares privatliv inte utspelar sig. Utmaningen skulle istället vara samarbetet stater emellan i och med att övervakningen i sådant fall hade skett på internationellt vatten (Rajamäki, 2018).

Flera artiklar som behandlar applikationer under tiden brottet begås föreslår testmodeller eller ramverk för ett visst system, utan någon större reflektion kring om det är möjligt att implementera med hänsyn till dagens teknik. Ett exempel på detta är artikeln skriven av Suralkar et al. (2020) som handlar om kontextmedveten AI. Idén om att AI i realtid skall kunna avgöra om en fysisk konfrontation mellan två människor är på allvar eller endast lek må vara god, men frågor uppstår kring hur realiserbart detta

faktiskt är. Det framgår exempelvis inte hur träningsdata för en sådan AI skulle se ut. Ett annat exempel är artikeln skriven av Enríquez et al. (2019), som föreslår realtidsrådgivning till användare vid situationer där ett potentiellt hot uppstår. Systemet förutsätter att alla personer inblandade har samma applikation installerad för att crowdsensingsystemet skall fungera, och det framgår inte hur systemet skall kunna uppfatta de nyanser som finns bland olika miljöer som rimligtvis bör ligga till grund för god rådgivning.

I avsnittet *Efter brott* utgörs en del av resultatet av möjligheten att se mönster och genomsöka data i syfte att upptäcka brott. En annan del kretsar kring bildanalys, som används med avsikten att ge spår vid brottsutredningar eller identifiera beteenden som tyder på att en person ljuger. Implementering av AI- eller algoritmbaserad bildanalys för syftet att ge spår vid brottsutredning kan tänkas vara mindre kontroversiellt än de tidigare nämnda användningsområdena för bildanalys, eftersom det är relativt likt de traditionella arbetsformer som används redan idag. Den gemensamma utgångspunkten är att det redan finns en misstanke om brott. Skillnaden blir att en dator går igenom videomaterialet i stället för en människa. Eftersom utgångspunkten är brottsmisstanke går det att hävda att en sådan implementering inte skulle omges av samma grad etisk risk som de som använder bildanalys före eller under tiden som brott begås.

Gemensamt för områdena *Före, Under,* och *Efter brott* är att författarna oftast inte kunnat uppge teknikens träffsäkerhet. Det leder fram till frågan hur potentiella felaktiga utslag bör hanteras vid en faktisk implementation. Vid tillfällen då felaktiga beslut har fattats av polismyndigheten har det hitintills inneburit att den arbetande polisen bär det eventuella straffansvaret för det begångna felet. Frågan blir om införandet av beslutsfattande teknik innebär ett överförande av ansvaret från människa till maskin. AI-utvecklade program innebär dessutom att programvaran lär sig själv baserat på träningsdata, vilket innebär att de beslut och antaganden som funktionaliteten baseras på inte alltid är explicit givna av en människa. Milivojevic (2021) lyfter därför frågan om vid vilken grad av självständighet AI-program bör behandlas som straffsskyldiga.

I litteratur utanför resultatet så rapporteras det att andra sektorer, såsom sjukvården och transporten, präglas av en liknande teknikutveckling men har skiftande ansvarsstrukturer. År 2011 införde Sverige en ny patientsäkerhetslag och blev då det första landet i världen som saknade individuellt ansvarsutkrävande av vårdpersonal för vårdskador (Brinkeback, 2018). Vidare genomgår fordonsindustrin en utveckling mot självkörande bilar, där besluten, istället för att fattas av föraren, fattas av algoritmer. I det fallet förespråkas det för att biltillverkaren bör bära ansvaret vid olyckor, inte föraren (Snowdon Smith, 2022).

Huruvida teknik eller människa kommer hållas ansvarig för felbedömningar kan också antas bero på hur människor, och samhället i stort, uppfattar tekniken. Enligt en studie utanför den undersökta litteraturen av Hong & Williams (2019) uppfattar människor brottsförutsägande AI som betydligt mindre självständiga än brottsförutsägande människor" Trots det visar studien att den allmänna uppfattningen hos folk är att en diskriminerande AI bär samma ansvar som en diskriminerande människa vid samma typ av prediktion. Denna tendens tyder på att en framtida utökning av AI inom

brottsbekämpning skulle kunna innebära ett minskat ansvarsutkrävande av personer, på grund av den ökade tilltron till tekniken.

Litteraturen förespråkar att hålla individer ansvariga för hur AI används, med motivationen att det ställer högre krav på kvaliteten hos programvaran, oavsett om det gäller användare eller utvecklare. Om individuellt ansvar utkrävs så försvinner möjligheten att gömma sig bakom algoritmerna när felaktiga beslut fattats. Det ställer i sin tur krav på att logiken bakom algoritmerna är begriplig och möjlig att uppfatta för människor. Den personliga integriteten och förtroendet för rättsprocessen skulle snabbt kunna urholkas ifall ett utpekande från ett AI-program som ingen förstod, eller var ansvarig för, räckte för att väcka misstanke eller åtal. Att hålla individer ansvariga ter sig alltså som en rimlig kompromiss, eftersom algoritmerna då kan användas utan att motivationerna bakom besluten som tas blir otillgängliga och obegripliga.

Det är alltså önskvärt med en transparent AI, inte bara sett till den tekniska implementeringen utan framför allt till förståelsen för utfallen den genererar. Intresset för utvecklandet av XAI-program är stort även i andra områden, men eftersom datan som hanteras är mycket känslig och konsekvenserna av misstag potentiellt blir enorma är det särskilt viktigt med sund logik bakom den automatiska beslutsfattningen i just brottsbekämpningen.

I litteraturen lyfts tendenser till att AI-programmen kan komma att agera diskriminerande mot vissa samhällsgrupper. Detta är knappast något som endast präglar brottsbekämpningen, då även källor utanför litteraturen rapporterat om implementationer av AI i andra områden som resulterat i olika former av diskriminering. Ett exempel är Microsoft Twitter-robot Tay, som efter endast en dags användning behövde tas ner på grund av dess rasistiska, antisemitiska och sexistiska uttalanden (Vincent, 2016). Ytterligare ett exempel är Amazons rekryteringsverktyg, som visade sig vara könsdiskriminerande mot kvinnor (Dastin, 2018). Det har varit först när AI-programmen väl använts i praktiken som problem som diskriminering och orättvisa visat sig, vilket tyder på vikten av praktisk tillämpning. Avsaknaden av praktisk implementation kan dock tänkas bero på det faktum att den träningsdata som krävs i brottsbekämpning ofta är känslig, och därmed omges av de integritetsproblem som lyfts i litteraturen. Även om så är fallet, talar det, om något, ännu mer för behovet av forskning i området.

Litteraturen lyfter fram risker associerade till massövervakning samt urholkning av privatliv och integritet. Utbredd övervakning på samhällsnivå kommer inte utan ovissheter. Några av artiklarna i litteraturen medger att det existerar en problematik gällande avvägningen mellan övervakning och personlig integritet, men besvarar inte det övertygande. I synnerhet inte när det kommer till artiklar som ser positivt på en ökad övervakning av olika delar i privatpersoners liv. Ofta försvaras det förhållandevis flyktigt med att det är för individens bästa eller att ändamålet helgar medlen.

Övervakning kan, förutom genom bildanalys, också åstadkommas genom text- och beteendeanalys. En vanlig plats för utvinning av sådan data är sociala medier. Under senare tid har det blivit alltmer välkänt att tillgång till sociala medier och appar i själva verket inte är gratis. I stället betalar användaren för sig genom att delge sin information

till plattformen. Ur den kontexten har koncensus blivit att användningen kommer ur det fria valet att använda tjänsten, men användaren kanske inte alltid är tillräckligt informerad om till vilken grad personlig information delas med tjänsteleverantörerna. Ändå råder en, om än tyst, enighet om att användaren har ett fritt val att ingå de avtal som användare och tjänsteleverantörer sluter, och att användarens samtycke legitimerar leverantörens informationsinsamling eftersom användaren, som en del av värdeutbytet, på så vis erhåller åtkomst till appen. Däremot tycks inte samma enighet gälla för polismyndighetens informationsinsamling. Om det beror på att medborgarna inte upplever att den information som de delar med polismyndigheten betalar av sig i tillräckligt många uppklarade brott för att helga medlen, eller individens egen rädsla att stå under ständig kontroll från rättsväsendets långa arm är ovisst.

För många kan det framstå som självklart att medborgarnas delning av personlig information bör vara frivillig, och inte påtvingad. Det beror på att det egna beslutet att dela informationen med myndigheter är avgörande för att myndighetsutövningen inte skall upplevas som en integritetskränkande aktivitet. Samtidigt inskränks medborgarnas rätt till att inte utsättas för brott i och med den överträdelse som brottslingen begår. Ur det uppstår frågan om vilken rättighet som är viktigast att prioritera, samt om en majoritet av medborgarna i ett samhälle är beredda att dela sin information med polismyndigheter för att proaktivt arbeta brottsförebyggande. Argumenten om integritetskränkning kan således komma från sidor som anser att den laglydige inte har någonting att dölja och därför borde vara bekväm med att dela sin information. Det bör däremot noteras att detta gäller under förutsättningen att de lagar som existerar också är fattade under demokratiska former, och stämmer överens med befolkningens uppfattning om vad som är rätt och fel. Det bör också noteras att lagar är av dynamisk karaktär. De förändras och byts ut i takt med att samhället utvecklas. Samhällsutvecklingen i sin tur drivs ofta av folkrörelser och engagemang. Med en omfattande övervakning kan en potentiell konsekvens bli att nytänkande och progressiva krafter i samhället hämmas.

Det verkar som frågan kokar ned i samhällets tolerans av viss reglementsvidrigt beteende från medborgare. En ständig informationsinsamling som berör hela samhället för polismyndighetens verksamhets skull kan skapa strukturer där personliga säregenheter tvingas bort, vilket gör samhället mer homogent samtidigt som det generella beteendet blir mer laglydigt. I lagen regleras olika brott med olika straffskala, helt enkelt för att brotten anses vara av olika grovhet. Kanske finns det på samma sätt motiv för att endast en viss sorts grövre brott motiverar insamling av personlig information från samhällets medborgare. På samma sätt bör i sådant fall inte personlig information få användas för upplösning av brott som inte anses vara av tillräckligt grov karaktär. Frågan handlar till slut om vilken grad av brottslighet som motiverar eventuell integritetskränkande informationsinsamling.

Interpols krav på *fairness, accountability, transparency,* och *explainability* är relevanta och kan användas som vägledning när det kommer till dessa etiska risker. Vid situationer där AI visar upp diskriminerande tendenser blir *fairness* förstås viktigt. *Accountability* kräver tydliga regelverk och riktlinjer för hur frågan om ansvar ska hanteras. *Transparency* och *explainability* har också mer eller mindre att göra med frågan om ansvar, och förhindrar att det uppstår en situation där man tvingas blint lita på

en algoritm som ingen förstår. Det finns inget krav som direkt behandlar integritetsproblem, men om de existerande kraven uppnås så kan det antas att övervakning, samt andra aktiviteter som kan ge upphov till etiska risker, genomförs på ett etiskt försvarbart vis.

Mot bakgrund av det resultat som presenterats och den diskussion som förts rekommenderas att mer empirisk forskning genomförs på området som kan visa upp konkreta metoder och tekniker som fungerar med den teknik tillgänglig idag. Förslagsvis bedrivs sådana projekt i samverkan med brottsbekämpande verksamheter för att bättre kunna pröva teknikens användbarhet i praktiken. Framtidens forskning bör även föra mer utvecklade resonemang kring de integritetsfrågor som uppstår i samband med insamling av den data mycket av den föreslagna tekniken är beroende av. I takt med att maskiners beslut får en allt större roll i samhället bör även mer forskning genomföras kring vem som bär ansvaret för dessa beslut, samt hur det kan försäkras att maskinernas beslut är rättvisa med hänsyn till mänskliga rättigheter.

5. Slutsats

Litteraturstudien visar att forskningsområdet *AI inom brottsbekämpning* fortfarande är i ett tidigt stadie. Optimismen för de nya digitala möjligheterna är stor men det är långt ifrån självklart hur dessa möjligheter skall vägas mot de etiska riskerna som medföljer dem. Litteraturens föreslagna lösningar förutsätter i regel stora datamängder i form av text-, bild- eller videomaterial. Insamlingen och hanteringen av detta material utgör inte bara etiska dilemman, utan även samordningsproblem. Vi föreslår därför att framtida studier inom området tar vid där dagens litteratur slutar, dels genom att fördjupa sig i de möjliga konsekvenserna av den föreslagna tekniken, dels genom att samverka med brottsbekämpande verksamhet och pröva teknikerna i verkligheten.

6. Reflektion om hållbarhet och etik

Som en del i analysen av arbetet ska ämnet ställas i relation till fem av FN:s Globala Mål. Nedan kommer de utvalda målen presenteras vartefter rapportens syfte analyseras utifrån målens definitioner.

De utvalda målen är:

- Mål 3: God hälsa och välbefinnande
- Mål 9: Hållbar industri, innovationer och infrastruktur
- Mål 11: Hållbara städer och samhällen
- Mål 16: Fredliga och inkluderande samhällen
- Mål 17: Genomförande och globalt partnerskap

Figur 3
Fem av FN:s globala mål

Kommentar. Ett urplock av FN:s globala mål som är del i Agenda 2030 för hållbar utveckling (Globala Målen, 2022)

6.1. Presentation av de fem utvalda globala målen

Först och främst är det viktigt att konstatera att en god hälsa och välbefinnande är grundläggande för en hållbar utveckling. I korthet går det att konstatera att medborgare med god hälsa har bättre förutsättningar att bidra till samhället och nå sin fulla potential. Främst fokus inom det här målet läggs på delmålen *3.5 förebygg och behandla drogmissbruk* samt *3.6 minska antalet dödsfall och skador i vägtrafiken* eftersom dessa händelser ofta är en följd av kriminella gärningar.

En annan dimension av hållbarhet som är värd att se till är hållbar industri, innovationer och infrastruktur som är viktiga aktörer inom hållbara, men också framgångsrika, samhällen. Främst ligger fokus på att skapa ekonomiskt och miljömässigt hållbara samhällen. Därför läggs fokus på delmål *9.1 skapa hållbara, motståndskraftiga och inkluderande infrastrukturer* som poängterar vikten av tillförlitlig och motståndskraftig infrastruktur.

Vidare behöver rollen av hållbara städer och samhällen presenteras som en del i ledet att skapa en hållbar framtid. I och med att städerna växer, växer även ekonomin och klassklyftorna kan bli större med risk för ökad social olydnad. Således ges delmål *11.3 inkluderande och hållbar urbanisering* speciellt fokus i och med delmålets specifika inriktning på att skapa välkomnande och trygga städer och stadsdelar.

Som fjärde mål fokuseras på fredliga och inkluderande samhällen med effektiva, pålitliga och ansvarsutkrävande institutioner. Med andra ord handlar det om att skapa trygga samhällen som är inkluderande och rättvisa. Särskilt koncentreras frågan till att handla om delmål *16.6 bygg effektiva, tillförlitliga och transparenta institutioner* som i sig handlar om att skapa institutioner som ges legitimitet av samhället. Samtidigt är det värt att poängtera att de resterande delmålen i mål 16 ligger som självklara latenta bikomponenter till målet.

Sist belyses målet om genomförande och globalt partnerskap som en självklarhet i och med globaliseringen av världen. Det handlar om att det i många fall går att lösa utmaningar genom att samla krafter och kompetens internationellt. Följaktligen är det delmål *17.17 uppmuntra effektiva partnerskap* som tar särskilt fokus på att underlätta för partnerskap mellan offentlig verksamhet, privat verksamhet och civilsamhället. De mellanstatliga samarbetet för utbyte av kompetens, information och andra erfarenheter ligger som en grundläggande pusselbit för att skapa hållbara samhällen världen över.

De ovan nämnda fem målen är, på ett eller annat sätt, kopplade till möjligheter och risker med att använda AI och algoritmer inom brottsbekämpande arbete. För att

förtydliga och nyansera kopplingarna följer en diskussion nedan som utgår från rapporten och de utvalda globala målen.

6.2. Analys av de fem utvalda globala målen

Rapporten har syftat till att kartlägga de möjligheter och risker som finns med AI och algoritmer i inom brottsbekämpning. Således bidrar rapporten till att sammanställa några förslag till åtgärder som bidrar till en mer hållbar värld genom de utvalda globala målen. Som resultatet i rapporten och presentationen av de utvalda globala målen gjort tydligt, finns starka kopplingar dem emellan. Nedan följer en redogörelse över deras skärningspunkter.

Det råder en uppfattning om att ju bättre polisen kan utföra sitt arbete, desto mer kommer god hälsa och välbefinnande främjas. Med breda penseldrag går det att finna synergier mellan tekniken och motarbetande av brott av diverse slag. Det ena av dessa omnämns i delmålen som beskriver hur ett förstärkt polisiärt arbete bidrar till ett bättre förebyggande arbetet mot drogmissbruk genom exempelvis övervakning av högriskindivider, videoanalys och språkanalys för beteendeprognostisering. Det andra beskriver hur antalet dödsfall och skador i vägtrafiken och det går att hämma genom videoanalys, både i statisk- och mobil form. Samtidigt bidrar dessa förebyggande och brottsbekämpande aktiviteter till det globala målet som berör fredliga och inkluderande samhällen.

Som tidigare nämnts består ett av delmålen som bygger upp det globala målet om fredliga och inkluderande samhällen om bygg effektiva, tillförlitliga och transparenta institutioner. Saken är den att polismyndighetens arbete med AI och realtidsanalys medför möjligheter som också stärker målet om tillförlitliga institutioner. Samtidigt måste då polismyndigheten upprätthålla tillräcklig nivå av transparens eftersom risken annars är stor att de trygga samhällen, som det globala målet om hållbara städer och samhällen berör, motarbetas. Mer konkret handlar frågan om de hållbara städerna i detta fallet om inkluderande och hållbar urbanisering, där myndigheternas verksamhet i närområdet blir avgörande för dess fortlevnad.

Fortlevnaden av polisväsendet i närområdet är grundläggande för att skapa och upprätthålla hållbar industri, innovationer och infrastruktur i ett samhälle. Utan polisens möjlighet att kontrollera medborgare fallerar samhällskontraktet och det mål som handlar om hållbara, motståndskraftiga och inkluderande infrastrukturer faller därmed. En fungerande infrastruktur är avgörande för att nå den sortens ekonomisk tillväxt som också omnämns som en del i de globala målen. Sammanfattningsvis verkar ändå slutsatsen få ligga inom målet om genomförande och globalt partnerskap.

Tillsammans blir människan starkare, och tillsammans blir även myndigheter starkare. Målet om genomförande och globalt partnerskap har ett gott initialt perspektiv, och det uppmuntrar till effektiva partnerskap. Frågan är bara om det finns en gräns där partnerskapen blir så starkt sammankopplade att de inte längre verkar för medborgarnas bästa.Samtidigt som det blir enklare om myndigheterna och länderna i världen blir mer sammanbundna, så finns en risk för att medborgarnas fri och rättigheter blir en konkurrensvara och till slut ett minne av en svunnen tid. Med det sagt är de hållbara

målen förenliga med resultatet i den här rapporten, men resultaten blir först bra i verkligheten när de implementeras av en godhjärtat människa.

Referenser

Anyoha, R. (2017). The History of Artificial Intelligence. *Science in the News Summer Edition 2017*. https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/

Behmer E.-J., Chandramouli K., Garrido V., Mühlenberg D., Müller D., Müller W., Pallmer D., Pérez F.J., Piatrik T., & Â Vargas C. (2019). Ontology Population Framework of MAGNETO for Instantiating Heterogeneous Forensic Data Modalities. Maglogiannis I.MacIntyre J.Iliadis L.Pimenidis E., *IFIP Advances in Information and Communication Technology* (s. 520-531). Springer New York LLC. 10.1007/978-3-030-19823-7_44

Brinkeback, D. (4 juni 2018). Det måste gå att utkräva ansvar i sjukvården. *Svenska Dagbladet*. https://www.svd.se/a/WL1Amj/det-maste-ga-att-utkrava-ansvar-i-sjukvarden

Bulgakova E., Bulgakov V., Trushchenkov I., Vasilev D., Kravets E. (2019). Big data in investigating and preventing crimes (181). *Springer International Publishing*. 10.1007/978-3-030-01358-5_6

Chang C.-Y., Chien L.-C., Kuo E.-C., & Hwan Y.-S. (2019). Designing Intelligence system of Image Processing and Mining in Cloud-Example of New Taipei City Police Department. Meen T.-H., *Proceedings of the 2nd IEEE International Conference on Knowledge Innovation and Invention 2019, ICKII 2019* (s. 293-295). Institute of Electrical and Electronics Engineers Inc.. 10.1109/ICKII46306.2019.9042641

Chase J., Du J., Fu N., Le T.V., & Lau H.C. (2018). Law enforcement resource optimization with response time guarantees. *2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings* (s. 1-7). Institute of Electrical and Electronics Engineers Inc.. 10.1109/SSCI.2017.8285326

Chen, L. Wong, G. (2019). Transcriptome Informatics. *Encyclopedia of Bioinformatics and Computational Biology Volume 2* (s. 324-340). 10.1016/B978-0-12-809633-8.20204-5

Contardo P., Sernani P., Falcionelli N., & Dragoni A.F. (2021). Deep learning for law enforcement: A survey about three application domains. *Xhina E.Hoxha K., CEUR Workshop Proceedings volume 2872* (s. 36-45). CEUR-WS.

Das P., & Das A.K. (2019). Application of Classification Techniques for Prediction and Analysis of Crime in India. Abraham A.Behera H.S.Naik B.Nayak J., *Advances in Intelligent Systems and Computing* (s. 191-201). Springer Verlag. 10.1007/978-981-10-8055-5_18

Dastin, J. (11 oktober 2018). *Amazon scraps secret AI recruiting tool that showed bias against women.* Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Denscombe, M. (2018). *Forskningshandboken* (4 uppl.). Studentlitteratur.

Du H., Xu Z., Yan Z., & Gao S. (2018). Intelligent video analysis technology of public security standard sets of data and measurements. Yen N.Y.Hung J.C.Hui L., *Lecture Notes in Electrical Engineering* (s. 453-456). Springer Verlag. 10.1007/978-981-10-7398-4_47

Enríquez F., Soria L.M., Álvarez-García J.A., Caparrini F.S., Velasco F., Deniz O., & Vallez N. (2019). Vision and crowdsensing technology for an optimal response in physical-security. Rodrigues J.M.F.Cardoso P.J.S.Monteiro J.Lam R.Krzhizhanovskaya V.V.Lees M.H.Sloot P.M.A.Dongarra J.J., *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (s. 15-26). Springer Verlag. 10.1007/978-3-030-22750-0_2

Frankenfield, J. (2021). *How Artificial Intelligence Works*. Investopedia. https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp.

Freeman S. (2020). Artificial intelligence for emergency management. Pham T.Solomon L.Rainey K., *Proceedings of SPIE - The International Society for Optical Engineering*. SPIE. 10.1117/12.2561636

Hong, J & Williams, D. (2019). Racism, responsibility and autonomy in HCI: Testing perceptions of an AI agent. *Computers in Human Behaviour*. 10.1016/j.chb.2019.06.012

Innefu. (u.å.). *How Artificial Intelligence In Policing Helps Crime Detection*. https://www.innefu.com/blog/how-artificial-intelligence-in-policing-helps-crime-detection

Interpol (2020). *Towards Responsible AI Innovation - Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement*. https://www.interpol.int/es/content/download/15290/file/AI%20Report%20INTERPOL%20UNICRI.pdf

Ionescu B., Ghenescu M., Rastoceanu F., Roman R., Buric M. (2020). Artificial Intelligence Fights Crime and Terrorism at a New Level. *IEEE Multimedia, 27(2)* (s. 55-61). 10.1109/MMUL.2020.2994403

Jie Y., Liu C.Z., Li M., Choo K.-K.R., Chen L., Guo C. (2020). Game theoretic resource allocation model for designing effective traffic safety solution against drunk driving. *Applied Mathematics and Computation, 376*. 10.1016/j.amc.2020.125142

Jindal S., & Sharma K. (2018). Intend to analyze Social Media feeds to detect behavioral trends of individuals to proactively act against Social Threats. Singh S.Asari V.K.Patel R.B.Sidike P., *Procedia Computer Science* (s. 218-225). Elsevier B.V.. 10.1016/j.procs.2018.05.191

Karolinska Institutet (2022). *Systematisk litteraturöversikt som examensarbete*. https://kib.ki.se/soka-vardera/systematiska-oversikter/systematisk-litteraturoversikt-som-examensarbete

Kaur B., Ahuja L., Kumar V. (2019). Decision tree Model: Predicting Sexual Offenders on the Basis of Minor and Major Victims. *Proceedings - 2019 Amity International Conference on Artificial Intelligence* (s. 193-197). 10.1109/AICAI.2019.8701276

Kitsos P. (2020). The limits of government surveillance: Law enforcement in the age of artificial intelligence. *CEUR Workshop Proceedings* (s. 164-168). CEUR-WS.

Lanagan S., Choo K.-K.R. (2021). On the need for AI to triage encrypted data containers in U.S. law enforcement applications. *Forensic Science International: Digital Investigation, 38*. 10.1016/j.fsidi.2021.301217

Matthew U.O., Kazaure J.S., Onyebuchi A., Daniel O.O., Muhammed I.H., & Okafor N.U. (2021). Artificial intelligence autonomous unmanned aerial vehicle (UAV) system for remote sensing in security surveillance., *Proceedings of the 2020 IEEE 2nd International Conference on Cyberspace, CYBER NIGERIA 2020* (s. 1-10). Institute of Electrical and Electronics Engineers Inc.. 10.1109/CYBERNIGERIA51635.2021.9428862

Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 1(2)*. 10.17169/fqs-1.2.1089

Milivojevic S. (2021). *Crime and punishment in the future internet: Digital frontier technologies and criminology in the twenty-first century.* Taylor and Francis. 10.4324/9781003031215

Molina-Molina J.C., Salhaoui M., Guerrero-González A., Arioua M. (2021). Autonomous marine robot based on ai recognition for permanent surveillance in marine protected areas. *Sensors, 21(8)*. 10.3390/s21082664

Nouri, S. (4 december 2020). How AI Is Making An Impact On The Surveillance World. *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2020/12/04/how-ai-is-making-an-impact-on-the-surveillance-world/

OHCHR. (u.å.). *Privacy in the digital age*. https://www.ohchr.org/en/privacy-in-the-digital-age

Pasquale, F. (2019). A rule of persons, not machines: The limits of legal automation. *Geo. Wash. L. Rev.*, *87*, 1.

Pawlicka A., Choraś M., Przybyszewski M., Belmon L., Kozik R., & Demestichas K. (2021). Why Do Law Enforcement Agencies Need AI for Analyzing Big Data?. Saeed K.Dvorsky J., *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (s. 331-342). Springer Science and Business Media Deutschland GmbH. 10.1007/978-3-030-84340-3_27

Rajamäki J., Sarlio-Siintola S., & Simola J. (2018). The ethics of open source intelligence applied by maritime law enforcement authorities. Josang A., *European Conference on Information Warfare and Security, ECCWS* (s. 424-431). Curran Associates Inc..

Rajapakshe C., Balasooriya S., Dayarathna H., Ranaweera N., Walgampaya N., & Pemadasa N. (2019). Using CNNs RNNs and Machine Learning Algorithms for Real-time Crime Prediction. *2019 International Conference on Advancements in Computing, ICAC 2019* (s. 310-316). Institute of Electrical and Electronics Engineers Inc.. 10.1109/ICAC49085.2019.9103425

Ramirez J., Campo-Archbold A., Zapata A., Diaz-Lopez D., Pastor-Galindo J., Gomez Marmol F., Aponte J. (2022). On the power of social networks to analyze threatening trends. *IEEE Internet Computing*. 10.1109/MIC.2022.3154712

Regeringskansliet. (16 maj 2018). *Nationell inriktning för artificiell intelligens*. https://www.regeringen.se/informationsmaterial/2018/05/nationell-inriktning-for-artificiell-intelligens/

Regeringskansliet. (u.å.). *Digitaliseringsstrategin*. https://www.regeringen.se/regeringens-politik/digitaliseringsstrategin/

Ridzuan Khairuddin A., Alwee R., & Haron H. (2020). A Comparative Analysis of Artificial Intelligence Techniques in Forecasting Violent Crime Rate. *IOP Conference Series: Materials Science and Engineering*. IOP Publishing Ltd. 10.1088/1757-899X/864/1/012056

Sherer J.A., Sterling N.L., Burger L., Banaschik M., Taal A. (2018). An investigator's christmas carol: Past, present, and future law enforcement agency data mining practices. *Cyber Criminology. Advanced Sciences and Technologies for Security Applications* (s. 251-273). Springer. 10.1007/978-3-319-97181-0_12

Simpson T. (2021). Real-Time Drone Surveillance System for Violent Crowd Behavior Unmanned Aircraft System (UAS) - Human Autonomy Teaming (HAT). *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*. Institute of Electrical and Electronics Engineers Inc.. 10.1109/DASC52595.2021.9594332

Singh A., Anand T., Sharma S., & Singh P. (2021). IoT Based Weapons Detection System for Surveillance and Security Using YOLOV4. *Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021* (s. 488-493). Institute of Electrical and Electronics Engineers Inc.. 10.1109/ICCES51350.2021.9489224

Smith M., Miller S. (2022). The ethical application of biometric facial recognition technology. *AI and Society, 37(1)* (s. 167-175). 10.1007/s00146-021-01199-9

Snowdon Smith, Z. (5 januari 2022). *Self-Driving Car Users Shouldn't Be Held Responsible For Crashes, U.K. Report Says.* Forbes. https://www.forbes.com/sites/zacharysmith/2022/01/25/self-driving-car-users-shouldnt-be-held-responsible-for-crashes-uk-report-says/?sh=1d99b5de37c9

Sun S. (2020). Application of fuzzy image restoration in criminal investigation. *Journal of Visual Communication and Image Representation*, *71*. 10.1016/j.jvcir.2019.102704

Suralkar S., Gangurde S., Chintakindi S., Chawla H. (2020). An Autonomous Intelligent Ornithopter (49). *Lecture Notes on Data Engineering and Communications Technologies* (s. 856-865). Springer Science and Business Media Deutschland GmbH. 10.1007/978-3-030-43192-1_93

Thoa Mac T., Copot C., Lin C.-Y., Hong Hai H., & Ionescu C.M. (2020). Towards the Development of a Smart Drone Police: Illustration in Traffic Speed Monitoring. *Journal of Physics: Conference Series.* Institute of Physics Publishing. 10.1088/1742-6596/1487/1/012029

Toppireddy H.K.R., Saini B., & Mahajan G. (2018). Crime Prediction & Monitoring Framework Based on Spatial Analysis. Singh S.Asari V.K.Patel R.B.Sidike P., *Procedia Computer Science* (s. 696-705). Elsevier B.V.. 10.1016/j.procs.2018.05.075

United Nations Development Programme (UNDP). Läs mer om ett av globala målen. Globala målen. https://www.globalamalen.se/

Vincent, J. *Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day.* The Verge. https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

Wang H., Ma S. (2022). Preventing crimes against public health with artificial intelligence and machine learning capabilities. *Socio-Economic Planning Sciences, 80*. 10.1016/j.seps.2021.101043

WASP- HS (u.å.). *About WASP-HS.* https://wasp-hs.org/about-wasp-hs/

Wyatt A. (2021). A Southeast Asian perspective on the impact of increasingly Autonomous systems on subnational relations of power. *Defence Studies, 21*(3) (s. 271-291). 10.1080/14702436.2021.1908136

Zhang R. (2021). The AI embedding predicts the legal risks of policing and its prevention. *Proceedings - 2021 International Conference on Computer Information Science and Artificial Intelligence, CISAI 2021* (s. 642-646). Institute of Electrical and Electronics Engineers Inc.. 10.1109/CISAI54367.2021.00129

INSTITUTIONEN FÖR TEKNIKENS EKONOMI OCH ORGANISATION
AVDELNINGEN FÖR TEKNIK, VETENSKAP OCH SAMHÄLLE
CHALMERS TEKNISKA HÖGSKOLA
Göteborg, Sverige 2022
www.chalmers.se

CHALMERS