

Cmd/Ctrl. Considerazioni etiche e operative sull'uso di sistemi di intelligenza artificiale in supporto alle decisioni militari

di Michele Carlo Tripeni *

L'uso crescente di sistemi di supporto decisionale basati sull'intelligenza artificiale nelle operazioni militari promette maggiore efficienza e rapidità, ma introduce il rischio dell'*automation bias*. Questo fenomeno può compromettere il giudizio umano e il rispetto del diritto internazionale umanitario. Il saggio esamina tali sfide e propone strategie per bilanciare l'uso dell'IA con il mantenimento del controllo umano.

Parole chiave: AI-DSS; automation bias; etica militare; processo decisionale; controllo umano; diritto internazionale umanitario.

Cmd/Ctrl. Ethical and operational considerations on the use of artificial intelligence systems to support military decisionmaking

The increasing use of Artificial Intelligence-based Decision Support Systems in military operations promises enhanced efficiency and speed but raises concerns about automation bias. This phenomenon may undermine human judgment and compliance with International Humanitarian Law. This paper explores these challenges and proposes strategies to balance AI advantages with the essential role of human oversight in military decision-making.

Keywords: AI-DSS; automation bias; military ethics; decision-making; meaningful human control; laws of war.

Introduzione

Con il crescente utilizzo nelle operazioni militari di sistemi di supporto decisionale basati sull'intelligenza artificiale (AI-DSS), vengono promessi miglioramenti senza precedenti in termini di efficienza, rapidità ed efficacia (tra gli altri: Cook, 2021; Harper, 2018). In effetti, questi sistemi sono in grado di setacciare enormi quantità di informazioni e di fornire raccomandazioni in tempo reale. Funzioni che attualmente sono essenziali per alleviare

DOI: 10.5281/zenodo.17524653

* University of Glasgow. michelec.tripeni@gmail.com.

il carico cognitivo del personale militare, il quale opera in contesti sempre più complessi e interconnessi. Tuttavia, questo utilizzo dell'intelligenza artificiale (IA) comporta anche notevoli sfide. Di particolare importanza nel contesto degli AI-DSS sono le problematiche legate all' *automation bias*. Ovvero, il fenomeno per cui gli operatori umani preferiscono deferire alle raccomandazioni generate dai sistemi automatizzati senza verificarle né interrogarle in modo critico. Tale dipendenza dall'IA nel contesto militare può introdurre vulnerabilità operative ed etiche di grande impatto, particolarmente per le loro conseguenze sul rispetto del diritto internazionale umanitario (DIU).

Una delle preoccupazioni più urgenti è proprio che il fenomeno, offuscando il giudizio etico umano degli operatori, possa portare a violazioni dei principi fondamentali del diritto bellico. Questi principi richiedono la valutazione di fattori morali ed operativi complessi da parte dei comandanti. Valutazione che anche i più avanzati sistemi di IA non sono in grado di compiere. In uno scenario di eccessiva dipendenza dalle raccomandazioni dell'IA, dunque, il rispetto delle norme di DIU e il controllo da parte dei comandanti militari potrebbero risultare sostanzialmente erosi. Portando così ad una situazione in cui gli operatori umani, e non tanto i "killer robot" al centro di molti dibattiti internazionali, vengono ridotti a fredde macchine assassine (Carpenter, 2024).

Questo saggio sostiene che il giudizio e la responsabilità finale devono risiedere negli operatori umani come garanzia indispensabile nelle applicazioni militari degli AI-DSS. Propone inoltre, sulla base della letteratura in materia, un approccio equilibrato che concili i vantaggi dell'IA in termini di efficienza ed efficacia con il mantenimento del controllo umano, fattore centrale per il comando militare e per la condotta legale della guerra. Il saggio fornisce inizialmente un breve resoconto dello sviluppo e dell'impiego di AI-DSS in ambito militare. Successivamente analizza la letteratura riguardo i fenomeni cognitivi legati al loro impiego, in particolare l'*automation bias*, e il loro potenziale impatto sul rispetto del DIU. Si conclude poi con alcuni suggerimenti per mitigare questo impatto e proteggere l'autonomia decisionale umana, oltre a fornire ulteriori spunti di riflessione per future ricerche nel campo.

1. Contesto e sviluppo degli AI-DSS nelle operazioni militari

I sistemi informatizzati per il supporto alle decisioni (*decision support systems*, DSS) non sono certamente una novità, né lo è l'idea di impiegare

tecniche di IA a questo scopo. Infatti, a partire almeno dalla metà degli anni Ottanta, nell'ambito del management e dell'informatica si è sviluppato il filone di ricerca sugli *intelligent decision support systems* (termine la cui genesi è attribuita a Holsapple e Whinston (1987)). Questi sistemi verranno successivamente ribattezzati *AI decision support systems* (AI-DSS) con l'affermarsi dell'intelligenza artificiale nel linguaggio comune. Già in questo contesto vi era la consapevolezza che per progettare al meglio tali sistemi fosse necessario comprendere gli utenti e le loro "limitazioni", oltre ai meccanismi secondo cui essi stessi determinassero quale metodo utilizzare per raggiungere una decisione. In questo senso l'obiettivo doveva essere presentare le informazioni in modo congruente con i processi decisionali degli utenti, aiutando semplicemente a scegliere lo strumento più adeguato e, al contempo, ad evitare errori nella manipolazione ed interpretazione dei dati. (Remus, Kottemann, 1986)

In ambito militare lo sviluppo degli AI-DSS ha una lunga storia parallela a quella delle applicazioni civili. I primi sistemi si focalizzavano su applicazioni di logistica nonché su automazioni di tipo *rule-based*, per poi includere sistemi IA sempre più avanzati andando a ricoprire funzioni di analisi d'intelligence e individuazione degli obiettivi. L'iniziativa lanciata nel 2017 dal Dipartimento della difesa americano, *Project Maven* (Manson, 2024), è l'esempio di questo cambiamento. Il sistema *Maven*, infatti, è progettato per utilizzare l'IA per processare grandi quantità di filmati da droni e satelliti, e assistere nell'identificazione dei bersagli. Un sistema simile è *Lavender*, il sistema sviluppato da Israele, che stando a quanto riportato da diverse testate è stato impiegato per generare enormi liste di obiettivi per i bombardamenti a Gaza (Abraham, 2024; Dvoskin, 2024; McKernan, Davies, 2024). Anche nella Guerra russo-ucraina vengono impiegati AI-DSS per il supporto a queste operazioni (Bondar, 2024). Tuttavia, il loro uso non è senza svantaggi. Molti ufficiali delle Forze armate statunitensi, ad esempio, hanno sollevato dubbi fondamentali sull'accuratezza e la robustezza di *Maven*. D'altro canto, l'uso di *Lavander* sembra aver favorito un approccio indiscriminato nella selezione degli obiettivi da parte di Israele, sollevando dubbi sulla sua affidabilità e sul rispetto del DIU. Da questi recenti esempi, sembra evidente il rischio di una dipendenza eccessiva dall'IA, particolarmente quando il controllo umano risulta diminuito, che può portare a complicazioni di tipo etico e legale.

2. Il rischio per il rispetto del diritto internazionale umanitario

La letteratura identifica come una delle preoccupazioni principali legate all'uso di AI-DSS quella dell'adesione alle norme di diritto umanitario internazionale (Nadibaidze *et al.*, 2024). In effetti, i sistemi di IA sono ampiamente ritenuti incapaci di interpretare correttamente i principi del DIU, in particolare quelli di proporzionalità e distinzione (Crootof, 2015; Sharkey, 2014; Szpak, 2020; Thurnher, 2018). Il principio di distinzione richiede di discriminare tra obiettivi di valore militare e non, valutando caso per caso. Il procedimento è ancora più complesso per bersagli umani (Crootof, 2015), in particolare in conflitti nei centri urbani, in cui la popolazione civile è più coinvolta e a volte partecipa direttamente alle ostilità, spesso senza un'evidente distinzione visiva tra combattenti e non (Szpak, 2020). Tuttavia, le tecnologie a disposizione delle forze armate ad oggi non permettono di compiere queste fini distinzioni.

In ogni caso, se pure la tecnologia avanzasse al punto da rendere possibile una corretta applicazione del principio di distinzione, il principio di proporzionalità sarebbe ancora un ostacolo. Il principio di proporzionalità, infatti, richiede di valutare caso per caso se il vantaggio di un'azione militare superi il danno collaterale previsto. Tuttavia, nonostante siano già in uso metodi computazionali per la valutazione della proporzionalità (ad esempio la *Collateral Damage Estimation Methodology*), questi sono pensati per fornire esclusivamente indicazioni di massima e non risultati definitivi. In più, intensificare il loro uso attraverso AI-DSS rischia di snaturare il principio di proporzionalità stesso (Gunnflo, Noll, 2023). D'altra parte, non esiste una regola fissa per confrontare il vantaggio militare di un'azione e i danni collaterali della stessa, né se ne dovrebbe derivare una da esempi passati, rendendo necessario valutare l'appropriatezza e legittimità di ciascun attacco caso per caso (Sharkey, 2014). Questo è completamente estraneo ai contemporanei sistemi di IA basati sul *machine learning* (ML), poiché essi basano la loro potenza sulla generalizzazione sulla base dei dati su cui sono addestrati.

Esistono anche problematiche di più ampio respiro. I sistemi ML sono noti per avere difficoltà nell'interpretazione dei principi giuridici in generale e nel "ragionamento giuridico" (Lai *et al.*, 2023). Inoltre, per interpretare accuratamente i principi di DIU è necessario una sorta di "giudizio umano" radicato nel buonsenso e nella buona fede (Szpak, 2020). Tuttavia, esiste un ampio consenso sulle limitate capacità di "ragionamento" dei sistemi ML (Bishop, 2021; Dentella *et al.*, 2024; Leivada *et al.*, 2024) e altrettanto sulla loro incapacità ad usare il "buonsenso" (Davis, Marcus, 2015).

3. L'automation bias e l'uso dell'IA in contesti militari

Come abbiamo visto sopra, nelle menti dei primi progettatori di AI-DSS vi era già la consapevolezza del rischio che potessero influenzare le decisioni umane ed introdurre in questo modo dei bias cognitivi. In effetti, questa si è confermata una delle problematiche principali dell'interazione uomo-macchina, compreso in campo militare (Dobbe, Wolters, 2024; Nadibaidze *et al.*, 2024). Alcuni dei primi studi in materia, effettuati nel campo dell'aviazione, hanno portato alla definizione del termine *automation bias* come gli “errori compiuti quando gli operatori umani utilizzano i suggerimenti automatizzati come un rimpiazzo euristico della raccolta ed analisi vigile delle informazioni” (Mosier *et al.*, 1998: 50-51). Tutt'oggi la maggior parte degli studi in materia si concentra in campo medico o aerospaziale (Goddard *et al.*, 2012), solo più recentemente l'impatto dell'*automation bias* è stato studiato anche nel campo della pubblica amministrazione (Ruscheimer, Hondrich, 2024; Schäferling, 2023), delle forze dell'ordine (Englezos, 2023; Selten *et al.*, 2023) e della giustizia (Gravett, 2024; Zerilli, 2022), mostrando come questo fenomeno si verifichi su ampia scala. Inoltre, l'*automation bias* può interagire con altri tipi di “scorciatoie mentali” come la *machine heuristic*, ovvero il meccanismo per cui gli umani sono portati a ritenere le macchine più “oggettive”, “imparziali” e “affidabili” degli umani (Sundar, 2008; Sundar, Kim, 2019). Il che può portare a una maggiore dipendenza dall'automazione e di conseguenza aumentare l'*automation bias* (Sundar, 2020).

Nel contesto militare, l'*automation bias*, in combinazione con i limiti degli AI-DSS esaminate nei primi due paragrafi, può avere conseguenze critiche come, ad esempio, la selezione errata dei bersagli, oltre a una generale riduzione della *situational awareness* e del “coinvolgimento morale” degli operatori. Ciò nonostante, lo studio di questo fenomeno in campo militare è ancora ai primordi, anche se vi è un'ampia evidenza proveniente da altri settori. Inoltre, malgrado la attuale mancanza di dati empirici sull'impatto dell'*automation bias* in ambito militare, non è difficile immaginare come questo possa svilupparsi. Infatti, secondo la *dual process theory* di Kahneman gli umani in condizioni di stress utilizzano i processi del “sistema 1”, più veloci e meno cognitivamente impegnativi, i quali tuttavia possono portare all'introduzione di bias cognitivi (Kahneman, 2011), come l'*automation bias*. Le operazioni militari sono un chiaro esempio di ambiente ad alto stress e gli ufficiali militari non sono meno suscettibili a questo genere di scorciatoie (Berejikian, Zwald, 2024; Knighton, 2004). Inoltre, sono già disponibili alcuni casi studio consolidati che dimostrano le implicazioni dell'*automation bias* in contesti operativi militari. Un caso emblematico riguarda gli incidenti

verificatisi durante l'impiego dei sistemi di difesa aerea *Patriot* nel corso dell'invasione dell'Iraq del 2003. In particolare, nel marzo di quell'anno, durante operazioni congiunte guidate dagli Stati Uniti, un sistema *Patriot* statunitense classificò erroneamente un velivolo militare britannico come un missile iracheno in arrivo. Dato il breve tempo per reagire a loro disposizione, gli operatori scelsero di aderire alle indicazioni fornite dal sistema, procedendo ad ingaggiare il velivolo. Pur avendo la possibilità di eseguire ulteriori verifiche e di disattendere l'allerta automatizzata, essi preferirono affidarsi al responso della macchina, abbattendo così erroneamente un velivolo amico. (Coco, 2024; Cummings, 2004; Stewart, Hinds, 2023). Alla luce di precedenti come questo, sembra quindi realistico il rischio che l'*automation bias* abbia un impatto significativo sulle decisioni dei comandanti militari, qualora utilizzino AI-DSS, con conseguenze potenzialmente catastrofiche sul rispetto delle norme di diritto bellico.

4. L'IA come strumento e non sostituto del comando umano

Per mitigare i rischi dell'uso di AI-DSS in ambito militare, dunque, è necessario assicurarsi un coinvolgimento umano continuativo e significativo durante tutto il ciclo di vita dei sistemi, dalla progettazione alla valutazione post-utilizzo. È necessario che gli operatori abbiano tutto il tempo e le risorse necessarie per valutare a pieno i risultati forniti dall'IA in combinazione con altre fonti. In particolare, se la velocità di elaborazione richiesta e il volume di informazioni sono eccessivi, gli operatori devono aver a disposizione chiari protocolli per compiere questa valutazione. È fondamentale riconoscere che accelerare il processo decisionale non è sempre vantaggioso e anzi può introdurre errori con conseguenze gravi, in particolare se l'ambiente operativo è complesso. Per questo bisogna stabilire limiti chiari alla velocità delle decisioni relative all'uso della forza, specialmente in contesti in cui esiste un alto rischio di danni collaterali (Nadibaidze *et al.*, 2024). Gli AI-DSS dovrebbero essere progettati per operare a una "velocità umana" anziché "velocità macchina", preservando così il ruolo del giudizio umano nelle decisioni sull'uso della forza (ICRC, 2020).

Per ridurre il rischio di *automation bias*, una strategia è l'uso di *confidence scores*, indicatori del livello di certezza di un output. Tuttavia, tali punteggi possono indurre una fiducia eccessiva negli operatori, che potrebbero non comprendere i limiti tecnici sottostanti. Per questo, le forze armate devono fornire linee guida che chiariscano che un alto livello di confidenza non equivale automaticamente alla legittimità di un bersaglio secondo il

Michele Carlo Tripeni

diritto internazionale (ICRC, 2024). Gli sviluppatori devono collaborare con i decisori militari per garantire un utilizzo informato di questi sistemi (Nadibaidze *et al.*, 2024.). Strumenti come le *model cards* possono fornire panoramiche sulle capacità e limitazioni di un sistema, aiutando gli operatori a prevedere possibili errori. Tuttavia, in ambito militare, la trasparenza di tali strumenti deve essere bilanciata con la necessità di evitare che le vulnerabilità dell'IA possano essere sfruttate da avversari (ICRC, 2024).

Infine, l'addestramento del personale militare deve includere una comprensione approfondita delle limitazioni sia umane che tecnologiche, con particolare attenzione ai problemi derivanti dall'interazione tra uomini e macchine in contesti operativi diversi. I programmi di formazione dovrebbero analizzare il modo in cui gli aspetti sociali, politici, istituzionali e tecnici interagiscono nell'uso degli AI-DSS, così da aggiornare le dottrine di *targeting* per rispondere alle sfide emergenti (Klaus, 2024; Nadibaidze *et al.*, 2024).

Conclusione e prospettive future

L'adozione degli AI-DSS nelle operazioni militari pone sfide significative, in particolare per il rischio di *automation bias* e le implicazioni sul rispetto del DIU. Sebbene questo articolo abbia esaminato tali questioni alla luce della letteratura esistente, resta la necessità di studi empirici che valutino concretamente l'impatto di questi sistemi nel contesto operativo. Un limite di questo lavoro è l'assenza di dati quantitativi e qualitativi relativi all'uso effettivo degli AI-DSS in scenari militari reali. La ricerca futura potrebbe concentrarsi su studi di casi dettagliati che analizzino incidenti operativi attribuibili all'*automation bias* o alla fiducia eccessiva nelle raccomandazioni dell'IA. Un'altra area di ricerca riguarda lo sviluppo e la valutazione di strategie di mitigazione, come l'uso di interfacce progettate per ridurre il bias o programmi di addestramento specifici. È inoltre cruciale indagare l'influenza di fattori organizzativi, culturali e istituzionali nella dipendenza dagli AI-DSS, poiché le dinamiche gerarchiche e dottrinali delle forze armate possono amplificare o attenuare il fenomeno. Infine, sarebbe utile un'analisi comparativa tra diversi contesti operativi e nazionali, per comprendere come variabili tecnologiche e normative influenzino l'integrazione dell'IA nelle decisioni militari. Solo attraverso un approccio empirico sistematico sarà possibile affinare le linee guida per un utilizzo etico e responsabile degli AI-DSS, garantendo che restino strumenti di supporto e non di deresponsabilizzazione.

Riferimenti bibliografici

- Abraham Y. (2024). 'Lavender': The AI machine directing Israel's bombing spree in Gaza. *+972 Magazine*. <https://www.972mag.com/lavender-ai-israeli-army-gaza/>
- Berejikian J.D., Zwald Z. (2024). Does military experience really matter? An empirical examination of decision framing and deterrence decision-making. *Journal of Global Security Studies*, 9(3). <https://doi.org/10.1093/jogss/ogae015>
- Bishop J.M. (2021). Artificial intelligence is stupid and causal reasoning will not fix it. *Frontiers in Psychology*, 11: 513474. <https://doi.org/10.3389/fpsyg.2020.513474>
- Bondar K. (2024). Understanding the military AI ecosystem of Ukraine. *Center for Strategic and International Studies*. <https://www.csis.org/analysis/understanding-military-ai-ecosystem-ukraine>
- Carpenter C. (2024). The real 'killer robots' are already here—and they're us. *World Politics Review*. <https://www.worldpoliticsreview.com/killer-robots-ai-israel-gaza/>
- Coco A. (2024). Exploring the impact of automation bias and complacency on individual criminal responsibility for war crimes. *Journal of International Criminal Justice*, 21(5): 1077-1096. <https://doi.org/10.1093/jicj/mqad034>
- Cook B. (2021). The future of artificial intelligence in ISR operations. *Air & Space Power Journal*, 35(Special Issue - Perspectives on JADO): 41-55.
- Crootof R. (2015). The killer robots are here: Legal and policy implications. *Cardozo Law Review*, 36(5): 1837-1915.
- Cummings M. (2004). Automation bias in intelligent time critical decision support systems. *ALAA 1st Intelligent Systems Technical Conference*: 1-6. <https://doi.org/10.2514/6.2004-6313>
- Davis E., Marcus G. (2015). Commonsense reasoning and commonsense knowledge in artificial intelligence. *Communications of the ACM*, 58(9): 92-103. <https://doi.org/10.1145/2701413>
- Dentella V., Günther F., Murphy E., Marcus G., Leivada E. (2024). Testing AI on language comprehension tasks reveals insensitivity to underlying meaning. *Scientific Reports*, 14(1): 28083. <https://doi.org/10.1038/s41598-024-79531-8>
- Dobbe R., Wolters A. (2024). Toward sociotechnical AI: Mapping vulnerabilities for machine learning in context. *Minds and Machines*, 34(2): 12. <https://doi.org/10.1007/s11023-024-09668-y>
- Dwoskin E. (2024). Israel built an 'AI factory' for war. It unleashed it in Gaza. *The Washington Post*. <https://www.washingtonpost.com/technology/2024/12/29/ai-israel-war-gaza-idf/>
- Englezos E. (2023). Policing by algorithm: NSW Police's suspect target management plan. *Alternative Law Journal*, 48(1): 17-24. <https://doi.org/10.1177/1037969X221147745>
- Goddard K., Roudsari A., Wyatt J.C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1): 121-127. <https://doi.org/10.1136/amiajnl-2011-000089>
- Gravett W.H. (2024). Judicial decision-making in the age of artificial intelligence. In Sousa Antunes H., Freitas P.M., Oliveira A.L., Martins Pereira C., Vaz De Sequeira E., Barreto Xavier L. (a cura di), *Multidisciplinary Perspectives on Artificial Intelligence and the Law*, vol. 58: 281-297. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-41264-6_15

Michele Carlo Tripeni

- Gunneflo M., Noll G. (2023). Technologies of decision support and proportionality in international humanitarian law. *Nordic Journal of International Law*, 92(1): 93-118. <https://doi.org/10.1163/15718107-bja10055>
- Harper J. (2018). Artificial intelligence to sort through ISR data glut. *National Defense*, 102(770): 33-35.
- Holsapple C.W., Whinston A.B. (1987). *Business expert systems*. Homewood (IL): Irwin.
- ICRC (2020). Artificial intelligence and machine learning in armed conflict: A human-centred approach. *International Review of the Red Cross*, 102(913): 463-479. <https://doi.org/10.1017/S1816383120000454>
- ICRC (2024). *Expert consultation report on AI and related technologies in military decision-making on the use of force in armed conflicts*. Geneva: International Committee of the Red Cross. <https://www.geneva-academy.ch/joomlatools-files/docman-files/Artificial%20Intelligence%20And%20Related%20Technologies%20In%20Military%20Decision-Making.pdf>
- Kahneman D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Klaus M. (2024). Transcending weapon systems: The ethical challenges of AI in military decision support systems. *Humanitarian Law & Policy*. <https://blogs.icrc.org/law-and-policy/2024/09/24/transcending-weapon-systems-the-ethical-challenges-of-ai-in-military-decision-support-systems/>
- Knighton R.J. (2004). The psychology of risk and its role in military decision-making. *Defence Studies*, 4(3): 309-334. <https://doi.org/10.1080/1470243042000344786>
- Lai J., Gan W., Wu J., Qi Z., Yu P.S. (2023). Large language models in law: A survey. *arXiv*. <https://doi.org/10.48550/arXiv.2312.03718>
- Leivada E., Marcus G., Günther F., Murphy E. (2024). A sentence is worth a thousand pictures: Can large language models understand human language and the world behind words? *arXiv*. <https://doi.org/10.48550/arXiv.2308.00109>
- Manson K. (2024). AI warfare is already here. *Bloomberg*. <https://www.bloomberg.com/features/2024-ai-warfare-project-maven/>
- McKernan B., Davies H. (2024). ‘The machine did it coldly’: Israel used AI to identify 37,000 Hamas targets. *The Guardian*. <https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes>
- Mosier K.L., Skitka L.J., Heers S., Burdick M. (1998). Automation bias: Decision making and performance in high-tech cockpits. *The International Journal of Aviation Psychology*, 8(1): 47-63. https://doi.org/10.1207/s15327108ijap0801_3
- Nadibaidze A., Bode I., Zhang Q. (2024). AI in military decision support systems: A review of developments and debates. *Center for War Studies*.
- Remus W.E., Kottemann J.E. (1986). Toward intelligent decision support systems: An artificially intelligent statistician. *MIS Quarterly*, 10(4): 403-418.
- Ruscheimer H., Hondrich L.J. (2024). Automation bias in public administration – an interdisciplinary perspective from law and psychology. *Government Information Quarterly*, 41(3): 101953. <https://doi.org/10.1016/j.giq.2024.101953>
- Schäferling S. (2023). Identifying challenges of governmental automated decision-making. In Schäferling S. (a cura di), *Governmental Automated Decision-Making and Human Rights*, vol. 62: 93-109. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-48125-3_4
- Selten F., Robeer M., Grimmelikhuijsen S. (2023). ‘Just like I thought’: Street-level bureaucrats trust AI recommendations if they confirm their professional judgment. *Public Administration Review*, 83(2): 263-278. <https://doi.org/10.1111/puar.13602>

Michele Carlo Tripeni

- Sharkey N. (2014). Towards a principle for the human supervisory control of robot weapons. *Politica & Società*, 2: 305-324. <https://doi.org/10.4476/77105>
- Stewart R., Hinds G. (2023). Algorithms of war: The use of artificial intelligence in decision making in armed conflict. *Humanitarian Law & Policy*. <https://blogs.icrc.org/law-and-policy/2023/10/24/algorithms-of-war-use-of-artificial-intelligence-decision-making-armed-conflict/>
- Sundar S.S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In Metzger M.J., Flanagin A.J. (a cura di), *Digital Media, Youth, and Credibility*: 73-100. Cambridge (MA): MIT Press. <https://doi.org/10.1162/dmal.9780262562324.073>
- Sundar S.S. (2020). Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). *Journal of Computer-Mediated Communication*, 25(1): 74-88. <https://doi.org/10.1093/jcmc/zmz026>
- Sundar S.S., Kim J. (2019). Machine heuristic: When we trust computers more than humans with our personal information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*: 1-9. <https://doi.org/10.1145/3290605.3300768>
- Szpak A. (2020). Legality of use and challenges of new technologies in warfare: The use of autonomous weapons in contemporary or future wars. *European Review*, 28(1): 118-131. <https://doi.org/10.1017/S1062798719000310>
- Thurnher J.S. (2018). Feasible precautions in attack and autonomous weapons. In Heintschel von Heinegg W., Frau R., Singer T. (a cura di), *Dehumanization of Warfare*: 99-117. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-67266-3_6
- Zerilli J. (2022). Algorithmic sentencing: Drawing lessons from human factors research. In Ryberg J., Roberts J.V. (a cura di), *Sentencing and Artificial Intelligence*: 165-183. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780197539538.003.0009>