Received: 22/11/2025
Review: 24/12/20265
Accepted: 04/02/2026
DOI: 10.22054/jocl.2025.8563.4282

Journal of Cyber Law
No(2),Vol(4), 129-142.
ISSN: 0972-6934
www.jocl.ir

Redefining National Security in the Age of Artificial Intelligence with a Focus on Technical, Legal, Ethical, and Governance Aspects

Ruhollah Bahreini [1*]
1*- PhD in Public Administration, Decision Making and Policy Making, Ct.C., Tehran, Iran.

ABSTRACT

The present study was conducted with the aim of examining various aspects of the use of artificial intelligence in the field of national security, using a descriptive review method. In this study, the existing literature on the applications of artificial intelligence in national security, the threats and challenges arising from it, the legal and ethical dimensions, as well as the related research gaps were reviewed and analyzed. The findings show that artificial intelligence has significant potential to improve the efficiency of security institutions in areas such as threat detection and prediction, large-scale security and intelligence data analysis, cyber defense, border monitoring, and countering psychological operations and disinformation. However, along with opportunities, this technology also brings risks such as privacy violations, algorithmic bias, ambiguity in legal liability, reduced human oversight, and the possibility of abuse by hostile actors. From a legal perspective, the need to develop clear frameworks for accountability, transparency, and oversight of intelligent systems became apparent. From an ethical perspective, principles such as justice, proportionality, human dignity, and harm minimization were emphasized. Also, the studies showed that existing research still faces serious gaps in integrating technical, legal, and ethical dimensions, as well as providing indigenous and practical models. Finally, this research emphasizes the need to develop responsible governance and smart regulation in the use of artificial intelligence in national security.

Keywords:
Artificial Intelligence, National Security, Security Threats, Cyber Defense

Article Type:Research Article

How to Cite: Bahreini,R . (2026). Redefining National Security in the Age of Artificial Intelligence with a Focus on Technical, Legal, Ethical, and Governance Aspects. *Journal of Cyber Law (JOCL)*, *2*(4), 129-142. doi: 10.22054/jocl.2025.8563.4282

Corresponding Author: Rozphd1402@gmail.com

نشریه حقوق سایبری
شماره (۱) ۱۴۲-۱۲۹
دوره ۱۴۰۴

تاریخ دریافت: ۱۴۰۴/۰۹/۰۱
تاریخ اصلاح ۱۴۰۴/۱۰/۰۳
تاریخ پذیرش: ۱۴۰۴/۱۱/۱۵

واکاوی بازتعریف امنیت ملی در عصر هوش مصنوعی با تمرکز بر ابعاد فنی، حقوقی، اخلاقی و حکمرانی

روح الله بحرینی[1*]

۱*- دکتری مدیریت دولتی، گرایش تصمیم‌گیری و خط و مشی گذاری، واحد تهران مرکزی، تهران، ایران.

چکیده

پژوهش حاضر با هدف بررسی ابعاد مختلف به‌کارگیری هوش مصنوعی در حوزه امنیت ملی، با روش مروری توصیفی انجام شده است. در این مطالعه، ادبیات موجود درباره کاربردهای هوش مصنوعی در امنیت ملی، تهدیدها و چالش‌های ناشی از آن، ابعاد حقوقی و اخلاقی، و نیز شکاف‌های پژوهشی مرتبط مورد بررسی و تحلیل قرار گرفت. یافته‌ها نشان می‌دهد که هوش مصنوعی در حوزه‌هایی مانند تشخیص و پیش‌بینی تهدیدها، تحلیل داده‌های امنیتی و اطلاعاتی در مقیاس بزرگ، دفاع سایبری، پایش مرزها و مقابله با عملیات روانی و اطلاعات نادرست، ظرفیت‌های قابل توجهی برای ارتقای کارآمدی نهادهای امنیتی دارد. با این حال، این فناوری در کنار فرصت‌ها، مخاطراتی همچون نقض حریم خصوصی، سوگیری الگوریتمی، ابهام در مسئولیت حقوقی، کاهش نظارت انسانی و احتمال سوءاستفاده توسط بازیگران متخاصم را نیز به همراه دارد. از منظر حقوقی، ضرورت تدوین چارچوب‌های روشن برای پاسخ‌گویی، شفافیت و نظارت بر سامانه‌های هوشمند آشکار شد. از منظر اخلاقی نیز رعایت اصولی مانند عدالت، تناسب، کرامت انسانی و حداقل‌سازی آسیب مورد تأکید قرار گرفت. همچنین، بررسی‌ها نشان داد که پژوهش‌های موجود هنوز در زمینه تلفیق ابعاد فنی، حقوقی و اخلاقی و نیز ارائه مدل‌های بومی و کاربردی، با شکاف‌های جدی مواجه‌اند. در نهایت، این پژوهش بر لزوم توسعه حکمرانی مسئولانه و تنظیم‌گری هوشمند در بهره‌گیری از هوش مصنوعی در امنیت ملی تأکید می‌کند.

کلیدواژه‌ها:

هوش مصنوعی، امنیت ملی، تهدیدهای امنیتی، دفاع سایبری

نوع مقاله: پژوهشی

نحوه استناد:

بحرینی، روح الله . (۱۴۰۴). بازتعریف امنیت ملی در عصر هوش مصنوعی با تمرکز بر ابعاد فنی، حقوقی، اخلاقی و حکمرانی. حقوق سایبری, ۲(۴), ۱۲۹-۱۴۲.

ایمیل نویسنده مسئول: Rozphd1402@gmail.com

مقدمه

در سال‌های اخیر، پیشرفت شتابان هوش مصنوعی به یکی از مهم‌ترین تحولات فناورانه جهان تبدیل شده و تأثیرات آن تنها به حوزه‌های صنعتی و اقتصادی محدود نمانده، بلکه به‌طور مستقیم بر مفاهیم بنیادینی مانند امنیت ملی نیز اثر گذاشته است. امنیت ملی در دنیای امروز دیگر صرفاً به توان نظامی یا مرزهای جغرافیایی خلاصه نمی‌شود، بلکه حوزه‌هایی چون امنیت سایبری، حفاظت از زیرساخت‌های حیاتی، مدیریت اطلاعات، مقابله با تهدیدهای نامتقارن، کنترل بحران‌ها و حتی جنگ روایت‌ها را نیز در بر می‌گیرد؛ حوزه‌هایی که همگی به‌نحوی با ظرفیت‌ها و مخاطرات هوش مصنوعی گره خورده‌اند. از یک سو، هوش مصنوعی می‌تواند با تحلیل سریع داده‌ها، شناسایی الگوهای تهدید، افزایش دقت تصمیم‌گیری و ارتقای توان دفاعی، به ابزاری مؤثر برای تقویت امنیت ملی تبدیل شود؛ و از سوی دیگر، همین فناوری می‌تواند زمینه‌ساز تهدیدهای تازه‌ای مانند حملات سایبری هوشمند، تولید و انتشار اطلاعات جعلی، نظارت گسترده، سلاح‌های خودمختار و افزایش آسیب‌پذیری نظام‌های حساس شود. بنابراین، بررسی نسبت میان هوش مصنوعی و امنیت ملی، نه‌تنها از منظر فناورانه بلکه از منظر سیاسی، حقوقی، اخلاقی و راهبردی نیز اهمیت فراوانی دارد.

در سال‌های اخیر، هوش مصنوعی به یکی از مهم‌ترین موضوعات در ادبیات امنیت ملی و امنیت سایبری تبدیل شده است، زیرا پژوهش‌های علمی جدید نشان می‌دهند این فناوری هم‌زمان می‌تواند نقش تقویت‌کننده و تهدیدآفرین داشته باشد؛ از یک سو، در مطالعات معتبر، بر ظرفیت هوش مصنوعی در تحلیل داده‌های حجیم، شناسایی الگوهای تهدید، ارتقای سامانه‌های تشخیص نفوذ، تقویت دفاع سایبری، پایش رفتار کاربران و بهبود واکنش در برابر حملات تأکید شده است (Kaur, Gabrijelcic, & Klobučar, 2023; Advancing cybersecurity and privacy with artificial intelligence, 2024; Mahmoudi & Bahrkazemi, 2024)، و از سوی دیگر، همین پژوهش‌ها هشدار می‌دهند که هوش مصنوعی می‌تواند برای خودکارسازی حملات سایبری، تولید فیشینگ و بدافزار، جعل هویت، دستکاری داده‌ها، افزایش نظارت و بازتولید سوگیری در فرایندهای امنیتی به کار گرفته شود (Taddeo, McCutcheon, & Floridi, 2019; Hobart, 2025; Artificial Intelligence and National Security, 2020). در ادبیات فارسی نیز پژوهش‌هایی با رویکرد توصیفی-تحلیلی نشان می‌دهند که هوش مصنوعی در حوزه‌هایی مانند امنیت عمومی، امنیت سایبری و حفاظت از داده‌ها، هم فرصت‌های قابل‌توجهی برای بهبود پیشگیری، پایش و تصمیم‌گیری فراهم می‌کند و هم چالش‌هایی مانند نقض حریم خصوصی، خطاهای الگوریتمی و دشواری نظارت انسانی را پیش روی سیاست‌گذاران می‌گذارد (بهرامی و ثابت‌سیار، ۱۴۰۳؛ محمودی و بحرکاظمی، ۱۴۰۳؛ رضایی و مرادی، ۱۴۰۳). بنابراین، آنچه از مرور این منابع برمی‌آید این است که بحث هوش مصنوعی در امنیت ملی دیگر صرفاً یک بحث فناورانه نیست، بلکه به مسئله‌ای چندبعدی در پیوند با حکمرانی، حقوق، اخلاق و راهبرد تبدیل شده است؛ مسئله‌ای که نیازمند مطالعه‌ای منظم، توصیفی و مبتنی بر شواهد علمی است تا هم ظرفیت‌های آن در تقویت امنیت روشن شود و هم ریسک‌های بالقوه‌اش به‌درستی فهم و مدیریت گردد.

در ادبیات جدید، هوش مصنوعی به‌عنوان یکی از مهم‌ترین فناوری‌های اثرگذار بر امنیت ملی، هم در سطح دفاعی و هم در سطح تهاجمی مورد توجه جدی پژوهشگران قرار گرفته است؛ به‌گونه‌ای که مطالعات اخیر نشان می‌دهند این فناوری از یک سو با توانایی تحلیل کلان‌داده‌ها، شناسایی الگوهای تهدید، افزایش سرعت واکنش در برابر حملات، و تقویت سازوکارهای امنیت سایبری، به ابزاری راهبردی برای دولت‌ها تبدیل شده است، و از سوی دیگر با کاهش موانع ورود به

حملات پیچیده، خودکارسازی عملیات مخرب، تولید جعل‌های عمیق، و ارتقای ظرفیت عملیات اطلاعاتی و روانی، سطح تازه‌ای از آسیب‌پذیری را پدید آورده است (Kaur, Gabrijelčič, & Klobućar, 2023; Taddeo, McCutcheon, & Floridi, 2019; Achuthan & Ramanathan, 2024). در همین راستا، پژوهش‌های به‌روزتر تأکید می‌کنند که تهدیدهای ناشی از AI دیگر صرفاً در قالب بدافزار یا نفوذ سایبری محدود نمی‌شوند، بلکه به حوزه‌هایی مانند دستکاری افکار عمومی، فیشینگ پیشرفته، حملات مبتنی بر مدل‌های زبانی بزرگ، و حتی تهدیدهای مرتبط با زیرساخت‌های حیاتی و عملیات نظامی نیز گسترش یافته‌اند؛ به همین دلیل، نهادهای سیاست‌گذار در سال‌های اخیر بیش از گذشته بر ضرورت حکمرانی ریسک‌محور، شفافیت الگوریتمی، نظارت انسانی، و چارچوب‌های ارزیابی و آزمون امنیتی تأکید کرده‌اند (National Security Memorandum on Artificial Intelligence, 2024; AI in Cybersecurity: Insights from the 2024 Benchmark Report, 2024). در ادبیات فارسی نیز مطالعات توصیفی–تحلیلی نشان می‌دهند که هوش مصنوعی در بستر امنیت ملی، هم فرصت‌هایی برای ارتقای قدرت بازدارندگی، تقویت دفاع سایبری و افزایش دقت تصمیم‌سازی فراهم می‌کند و هم با ایجاد امکان نظارت گسترده، نقض حریم خصوصی، و برجسته‌سازی سوگیری‌های الگوریتمی، چالش‌های حقوقی و اخلاقی مهمی را پیش روی دولت‌ها قرار می‌دهد (دهقانی فیروزآبادی و چهرزاد، ۱۴۰۳؛ احمدی و همکاران، ۱۴۰۱). بنابراین، جمع‌بندی پژوهش‌های جدید حاکی از آن است که هوش مصنوعی را باید نه صرفاً یک ابزار فناورانه، بلکه یک متغیر راهبردی در بازتعریف معماری امنیت ملی دانست؛ متغیری که اگرچه می‌تواند ظرفیت‌های دفاعی و بازدارنده دولت‌ها را افزایش دهد، اما در صورت فقدان تنظیم‌گری مؤثر، به تشدید رقابت‌های امنیتی، افزایش نابرابری قدرت و تضعیف برخی حقوق بنیادین نیز منجر خواهد شد.

کاربردهای هوش مصنوعی در امنیت ملی

تشخیص و پیش‌بینی تهدیدها

تشخیص و پیش‌بینی تهدیدها یکی از مهم‌ترین کاربردهای هوش مصنوعی در امنیت ملی است، زیرا این فناوری با توانایی پردازش حجم عظیمی از داده‌ها در زمان کوتاه می‌تواند الگوهای پنهان، نشانه‌های اولیه بحران، رفتارهای غیرعادی و روندهای تهدیدزا را قبل از تبدیل‌شدن به یک خطر جدی شناسایی کند. در واقع، سامانه‌های مبتنی بر هوش مصنوعی با تحلیل داده‌های اطلاعاتی، شبکه‌های ارتباطی، تصاویر ماهواره‌ای، داده‌های سایبری و حتی رفتار کاربران، امکان پیش‌بینی حملات، ارزیابی سطح ریسک و هشدار زودهنگام را فراهم می‌کنند؛ قابلیتی که در شرایط پیچیده امنیتی، می‌تواند به تصمیم‌گیری سریع‌تر و دقیق‌تر نهادهای مسئول کمک کند (Kaur et al., 2023; Taddeo et al., 2019; Achuthan et al., 2024). از سوی دیگر، پژوهش‌های جدید نشان می‌دهند که این توان پیش‌بینی‌گر اگرچه قدرت دفاعی دولت‌ها را افزایش می‌دهد، اما در صورت وابستگی بیش از حد به الگوریتم‌ها ممکن است به خطاهای تحلیلی، سوگیری داده‌ای و تصمیم‌های نادرست نیز منجر شود؛ بنابراین استفاده از هوش مصنوعی در این حوزه باید همراه با نظارت انسانی و چارچوب‌های کنترلی دقیق باشد (Kaur et al., 2023; Dehghani Firoozabadi & Chehrazad, 2023).

تحلیل داده‌های امنیتی و اطلاعاتی در مقیاس بزرگ

تحلیل داده‌های امنیتی و اطلاعاتی در مقیاس بزرگ یکی از مهم‌ترین ظرفیت‌های هوش مصنوعی در حوزه امنیت ملی است، زیرا حجم عظیم داده‌های تولیدشده از منابعی مانند سامانه‌های نظارتی، پایش شبکه‌های اجتماعی، داده‌های سایبری، تصاویر ماهواره‌ای و گزارش‌های اطلاعاتی، فراتر از توان تحلیل سنتی است. در چنین شرایطی، هوش مصنوعی می‌تواند با شناسایی الگوهای پنهان، یکپارچه‌سازی داده‌های پراکنده و استخراج نشانه‌های معنادار از میان اطلاعات خام، به نهادهای امنیتی در تشخیص تهدیدها و تصمیم‌سازی دقیق‌تر کمک کند (Dai & Boroomand, 2021; Dilmaghani et al., 2019). همچنین، این فناوری با افزایش سرعت پردازش و ارتقای دقت تحلیل، امکان پایش مداوم محیط‌های پرریسک و شناسایی رفتارهای غیرعادی را فراهم می‌سازد؛ قابلیتی که در شرایط پیچیده امنیتی، ارزش راهبردی بالایی دارد (National Security Agency et al., 2025; Kalpande et al., 2025). با این حال، موفقیت این فرایند به کیفیت داده‌ها، امنیت سامانه‌ها و نظارت انسانی وابسته است، زیرا استفاده نادرست از مدل‌های هوش مصنوعی می‌تواند به خطاهای تحلیلی و تصمیم‌های پرهزینه منجر شود (Sangarsu, 2018).

دفاع سایبری و شناسایی حملات

دفاع سایبری و شناسایی حملات از مهم‌ترین کاربردهای هوش مصنوعی در امنیت ملی است، زیرا زیرساخت‌های حیاتی، شبکه‌های ارتباطی، سامانه‌های مالی و پایگاه‌های داده به‌طور مداوم در معرض تهدیدهایی مانند بدافزارها، فیشینگ، نفوذهای پیشرفته، باج‌افزارها و حملات توزیع‌شده محروم‌سازی از خدمت قرار دارند. در این میان، هوش مصنوعی با تحلیل لحظه‌ای ترافیک شبکه، شناسایی الگوهای غیرعادی، تشخیص رفتارهای مشکوک و تطبیق داده‌های ورودی با الگوهای حملات شناخته‌شده، می‌تواند تهدیدها را بسیار سریع‌تر از روش‌های سنتی شناسایی کند و حتی پیش از گسترش حمله، هشدار لازم را صادر نماید (NIST, 2024; ENISA, 2023). همچنین، سامانه‌های مبتنی بر یادگیری ماشین قادرند با یادگیری از داده‌های گذشته، حملات پیچیده و ناشناخته را نیز از طریق شاخص‌های رفتاری و آماری تشخیص دهند و در نتیجه، قدرت واکنش تیم‌های امنیتی را افزایش دهند (Buczak & Guven, 2016; Sarker et al., 2021). علاوه بر این، هوش مصنوعی در خودکارسازی پاسخ به رخدادها، اولویت‌بندی هشدارها و کاهش بار کاری تحلیل‌گران امنیتی نقش مهمی دارد؛ موضوعی که به‌ویژه در سازمان‌های بزرگ و زیرساخت‌های حساس، به بهبود سرعت واکنش و کاهش خسارت کمک می‌کند (IBM Security, 2024). با وجود این مزایا، اثربخشی این فناوری به کیفیت داده‌ها، به‌روزرسانی مداوم مدل‌ها و نظارت انسانی وابسته است، زیرا مهاجمان نیز می‌توانند با حملات فریب‌دهنده یا دستکاری داده‌ها، مدل‌های هوشمند را دچار خطا کنند (Buczak & Guven, 2016).

پایش مرزها و شناسایی فعالیت‌های مشکوک

پایش مرزها و شناسایی فعالیت‌های مشکوک یکی از کاربردهای مهم هوش مصنوعی در امنیت ملی است، زیرا مرزهای زمینی، دریایی و هوایی معمولاً در معرض طیف گسترده‌ای از تهدیدها مانند قاچاق، عبور غیرقانونی، نفوذ افراد ناشناس، جابه‌جایی تجهیزات مشکوک و فعالیت‌های خرابکارانه قرار دارند. در این زمینه، هوش مصنوعی با استفاده از تحلیل تصاویر دوربین‌های نظارتی، داده‌های حسگرها، پهپادها، سامانه‌های تشخیص حرکت و اطلاعات مکانی می‌تواند رفتارهای غیرعادی را به‌صورت لحظه‌ای شناسایی کرده و میان الگوهای عادی و غیرعادی تمایز قائل شود (Hammoudeh et

al., 2020; Meunier et al., 2021). همچنین، این فناوری با ترکیب داده‌های چندمنبعی و بهره‌گیری از یادگیری ماشین، امکان تشخیص سریع‌تر تهدیدهای بالقوه و کاهش خطای انسانی را فراهم می‌کند؛ به‌ویژه در مناطق وسیع یا کم‌دسترس که پایش سنتی آن‌ها بسیار پرهزینه و زمان‌بر است (Agarwal et al., 2022; Chen et al., 2023). با این حال، کارایی این سامانه‌ها به کیفیت داده‌های ورودی، دقت مدل‌های تشخیص و هماهنگی آن‌ها با نیروهای انسانی وابسته است، زیرا شرایط محیطی، خطای سنسورها و رفتارهای فریب‌دهنده می‌توانند بر نتیجه تحلیل اثر بگذارند (Hammoudeh et al., 2020).

مقابله با عملیات روانی، جعل عمیق و اطلاعات نادرست

مقابله با عملیات روانی، جعل عمیق و اطلاعات نادرست یکی از چالش‌های جدی امنیت ملی در عصر دیجیتال است، زیرا این نوع تهدیدها با بهره‌گیری از شبکه‌های اجتماعی، رسانه‌های جعلی، محتوای دست‌کاری‌شده و فناوری‌های تولید متن، تصویر و ویدئوی مصنوعی می‌توانند افکار عمومی را دچار تردید، بی‌اعتمادی و آشفتگی کنند. در این میان، هوش مصنوعی با تحلیل الگوهای انتشار محتوا، شناسایی حساب‌های غیرواقعی، تشخیص هماهنگی‌های مشکوک در شبکه‌ها و بررسی ویژگی‌های فنی محتوای چندرسانه‌ای، می‌تواند به کشف سریع‌تر عملیات اطلاع‌رسانی هدفمند و جعل‌های عمیق کمک کند (Chesney & Citron, 2019; Tolosana et al., 2020). همچنین، مدل‌های یادگیری ماشین قادرند با تحلیل نشانه‌های زبانی، تصویری و رفتاری، محتوای نادرست را از اطلاعات معتبر تفکیک کرده و به نهادهای مسئول در اولویت‌بندی و پاسخ‌گویی مؤثرتر یاری رسانند (Vaccari & Chadwick, 2020; West et al., 2019). با این حال، موفقیت این رویکردها وابسته به به‌روزرسانی مداوم مدل‌ها، دسترسی به داده‌های دقیق و نظارت انسانی است، زیرا مهاجمان نیز به‌سرعت روش‌های فریب و تولید محتوای جعلی را ارتقا می‌دهند (Chesney & Citron, 2019).

تهدیدها و چالش‌ها

هوش مصنوعی در امنیت ملی، هم یک ابزار توانمندساز و هم یک منشأ تهدیدهای تازه است؛ از یک سو می‌تواند سرعت تحلیل، کشف تهدید و تصمیم‌سازی را افزایش دهد، اما از سوی دیگر، در صورت سوءاستفاده یا طراحی نامناسب، به‌عنوان عاملی برای تشدید ریسک‌های امنیتی عمل می‌کند (Horowitz & Kahn, 2024). یکی از مهم‌ترین چالش‌ها، بهره‌برداری مهاجمان از هوش مصنوعی برای تولید جعل عمیق، عملیات فریب و دست‌کاری اطلاعاتی است؛ به‌طوری‌که محتوای مصنوعی می‌تواند اعتماد عمومی به رسانه‌ها و نهادهای رسمی را تضعیف کرده و در بحران‌ها موجب سردرگمی، بی‌ثباتی و اختلال در واکنش‌های امنیتی شود (Liu et al., 2024; U.S. Department of Defense, 2023). در کنار این، هوش مصنوعی می‌تواند حملات سایبری را نیز هوشمندتر، سریع‌تر و پنهان‌تر کند؛ از جمله با امکان تولید فیشینگ‌های متقاعدکننده، بدافزارهای تطبیقی و حملات خودکار علیه سامانه‌های حساس، که سطح تهدید برای زیرساخت‌های حیاتی را به‌طور محسوسی بالا می‌برد (Yoganathan et al., 2025). از منظر درونی نیز، سوگیری الگوریتمی و اتکای بیش از حد به تصمیم‌های خودکار می‌تواند در حوزه‌های حساس امنیتی به خطاهای پرهزینه منجر شود؛ برای نمونه، پژوهش‌ها نشان می‌دهند که automation bias ممکن است باعث شود تصمیم‌گیران انسانی بیش از اندازه به خروجی سامانه اعتماد کنند، حتی وقتی زمینه تصمیم‌گیری پیچیده و پرریسک است (Horowitz & Kahn, 2024). همچنین، آسیب‌پذیری مدل‌ها در برابر prompt injection، model poisoning و داده‌های آلوده می‌تواند موجب افشای اطلاعات، انحراف

رفتار مدل و کاهش قابلیت اعتماد سامانه‌های امنیتی شود؛ به‌ویژه در سامانه‌های مبتنی بر LLM که پژوهش‌های جدید نشان داده‌اند در برابر این نوع حملات بسیار شکننده‌اند (Zou et al., 2024; Guo & Cai, 2025). در نتیجه، استفاده از هوش مصنوعی در امنیت ملی بدون نظارت انسانی، ارزیابی مستمر، کنترل دسترسی، و سازوکارهای حکمرانی شفاف، می‌تواند به‌جای تقویت امنیت، خود به یک سطح جدید از تهدید تبدیل شود (Horowitz & Kahn, 2024; U.S. Department of Defense, 2023).

ابعاد حقوقی

از منظر حقوقی، استفاده از هوش مصنوعی در حوزه امنیت ملی با چالش‌های جدی در زمینه حریم خصوصی، مسئولیت مدنی، شفافیت، نظارت‌پذیری و انطباق با حقوق بنیادین همراه است؛ زیرا سامانه‌های هوشمند در فرآیندهای پایش، تحلیل داده و تصمیم‌سازی امنیتی می‌توانند به جمع‌آوری گسترده داده‌های شخصی، نظارت فراگیر و حتی نقض حق حریم خصوصی منجر شوند و در نتیجه ضرورت محدودسازی مبتنی بر قانون، تناسب و ضرورت را برجسته می‌سازند (Babuta et al., 2020; Gollatz et al., 2018). افزون بر این، هنگامی که تصمیم یا اقدام زیان‌بار از سوی یک سامانه هوش مصنوعی رخ می‌دهد، تعیین مسئولیت حقوقی میان طراح، تولیدکننده، بهره‌بردار و نهاد ناظر دشوار می‌شود و همین امر موجب شده است که پژوهش‌های حقوقی، مسئولیت مبتنی بر تقصیر، مسئولیت محض و مسئولیت تضامنی را به‌عنوان الگوهای قابل بررسی مطرح کنند (Gless & Bertolini, 2022; Zakerinia, 2023). در ادبیات فارسی نیز تأکید شده است که هوش مصنوعی فاقد شخصیت حقوقی مستقل و اراده انسانی است و بنابراین نمی‌توان مسئولیت را به خودِ سامانه منتسب کرد؛ بلکه مسئولیت باید بر اساس عرف، تسبیب و میزان تأثیر میان عوامل انسانی مرتبط توزیع شود (Bagheri, n.d.; Hekmatnia et al., 2019; Alipour et al., 2024). از سوی دیگر، استفاده امنیتی از هوش مصنوعی می‌تواند با سوگیری الگوریتمی، تبعیض، و خطاهای تصمیم‌گیری همراه شود و این امر حق دادرسی منصفانه، حق برابری و حتی آزادی بیان را در معرض تهدید قرار می‌دهد؛ به‌ویژه در حوزه‌هایی مانند نظارت امنیتی، پیش‌بینی خطر و کنترل مرزی که آثار آن مستقیماً بر حقوق شهروندان می‌نشیند (Access Now, 2018; Council of Europe, 2018; Crawford, 2019). در نتیجه، بسیاری از پژوهش‌ها بر این نکته تأکید دارند که تنظیم‌گری مؤثر هوش مصنوعی در امنیت ملی باید بر پایه شفافیت الگوریتمی، نظارت انسانی، پاسخ‌گویی حقوقی و سازوکار جبران خسارت استوار باشد تا توازن میان امنیت عمومی و حقوق اساسی افراد حفظ شود (Babuta et al., 2020; Zakerinia & Gholampour, 2024; Hosseini et al., 2024).

ابعاد اخلاقی

از منظر اخلاقی، به‌کارگیری هوش مصنوعی در حوزه امنیت ملی با پرسش‌های بنیادینی درباره عدالت، کرامت انسانی، شفافیت، پاسخ‌گویی و میزان مجاز مداخله در زندگی افراد همراه است؛ زیرا سامانه‌های هوشمند ممکن است در فرآیندهای نظارتی، پیش‌بینی تهدید و تصمیم‌سازی امنیتی، منجر به کاهش اختیار انسانی، تقویت تبعیض‌های موجود و عادی‌سازی نظارت گسترده شوند، به‌ویژه اگر داده‌های آموزشی سوگیرانه باشند یا سازوکارهای کنترلی کافی وجود نداشته باشد (Mittelstadt et al., 2016; Selbst et al., 2019). از سوی دیگر، یکی از نگرانی‌های اخلاقی مهم این است که تصمیم‌های مبتنی بر الگوریتم، به‌ظاهر عینی و بی‌طرف جلوه کنند، در حالی که در عمل می‌توانند بازتاب‌دهنده ارزش‌ها،

اولویت‌ها و سوگیری‌های پنهان طراحان و نهادهای بهره‌بردار باشند؛ به همین دلیل، مسئله «پاسخ‌گویی اخلاقی» در کنار پاسخ‌گویی فنی اهمیت ویژه‌ای پیدا می‌کند (Floridi et al., 2018; Crawford, 2021). علاوه بر این، در امنیت ملی ممکن است استدلال شود که منافع جمعی ایجاب می‌کند از ابزارهای پرقدرت هوش مصنوعی استفاده شود، اما از منظر اخلاق کاربردی، این توجیه تنها زمانی پذیرفتنی است که اصل تناسب، ضرورت، حداقل‌سازی آسیب و نظارت انسانی واقعی رعایت شود؛ در غیر این صورت، امنیت بهانه‌ای برای تضعیف کرامت و حقوق افراد خواهد شد (Russell, 2019; Jobin et al., 2019). همچنین، استفاده از هوش مصنوعی برای شناسایی چهره، امتیازدهی ریسک یا پیش‌بینی رفتار، می‌تواند افراد را به «ابژه‌های داده‌ای» تقلیل دهد و از منظر اخلاقی مسئله‌ساز باشد، زیرا انسان را نه به‌عنوان فاعل اخلاقی، بلکه به‌عنوان مجموعه‌ای از الگوهای قابل‌محاسبه بازنمایی می‌کند (Crawford, 2021; Bostrom, 2014). در نتیجه، ادبیات اخلاق هوش مصنوعی تأکید می‌کند که توسعه و استفاده از این فناوری در امنیت ملی باید با شفافیت، ممیزی اخلاقی، قابلیت توضیح‌پذیری، نظارت انسانی و سازوکارهای پیشگیری از آسیب همراه باشد تا کارآمدی امنیتی به قیمت فرسایش اعتماد عمومی و ارزش‌های انسانی تمام نشود (Floridi et al., 2018; Jobin et al., 2019; Russell, 2019).

شکاف‌های پژوهشی

با توجه به پژوهش‌های داخلی و خارجی، مهم‌ترین شکاف پژوهشی در این حوزه آن است که اغلب مطالعات یا بر جنبه‌های فنی و کارکردی هوش مصنوعی در امنیت ملی تمرکز کرده‌اند یا صرفاً به‌صورت هنجاری از حقوق، اخلاق و حکمرانی سخن گفته‌اند، اما کمتر پژوهشی توانسته این دو سطح را به‌طور یکپارچه و در قالب یک مدل تحلیلی منسجم به هم پیوند دهد؛ به‌ویژه در مورد اینکه یک سامانه هوشمند در مراحل مختلف چرخه عمر خود، از طراحی و آموزش تا استقرار و نظارت، چه آثار حقوقی و اخلاقی مشخصی ایجاد می‌کند (Babuta et al., 2020; Mittelstadt et al., 2016). در ادبیات خارجی، خلأ مهم دیگر در عملیاتی‌سازی اصولی مانند شفافیت، پاسخ‌گویی، تناسب و نظارت انسانی است؛ یعنی بسیاری از آثار این اصول را تأیید می‌کنند، اما هنوز چارچوب‌های آزمون‌پذیر و اجرایی برای ممیزی الگوریتمی، ثبت تصمیم، تخصیص مسئولیت و امکان اعتراض مؤثر ارائه نشده است (Floridi et al., 2018; Gless & Bertolini, 2022; Selbst et al., 2019). در پژوهش‌های داخلی نیز هرچند درباره مسئولیت مدنی، مبانی فقهی و چالش‌های حقوقی هوش مصنوعی مقالات ارزشمندی نوشته شده است، اما این آثار غالباً توصیفی یا نظری‌اند و کمتر به مطالعه موردی، داده‌های تجربی یا مقایسه تطبیقی با نظام‌های حقوقی دیگر پرداخته‌اند؛ در نتیجه هنوز روشن نیست که در بستر ایران، کدام خطرها برجسته‌ترند: تبعیض الگوریتمی، نقض حریم خصوصی، ابهام مسئولیت یا فاصله میان سیاست‌گذاری و اجرا (Hekmatnia et al., 2019; Alipour et al., 2024; Zakerinia & Gholampour, 2024). بنابراین، پژوهش‌های آینده باید از سطح کلی‌گویی عبور کرده و با رویکردی میان‌رشته‌ای، مبتنی بر مطالعه موردی و تطبیقی، نشان دهند که چگونه می‌توان هم کارآمدی امنیتی را حفظ کرد و هم از نظارت فراگیر، تصمیم‌گیری غیرپاسخ‌گو و فرسایش حقوق بنیادین جلوگیری نمود (Crawford, 2021; Jobin et al., 2019).

پیشنهاد برای پژوهشگران آتی:

۱. مطالعه موردی یک کاربرد مشخص هوش مصنوعی در امنیت ملی.

۲. پژوهش تطبیقی میان ایران، اتحادیه اروپا و آمریکا درباره نظارت و مسئولیت.

۳. طراحی مدل بومی برای ارزیابی تأثیر حقوقی و اخلاقی سامانه‌های امنیتی.

۴. بررسی تجربی سوگیری الگوریتمی در داده‌ها و تصمیم‌های امنیتی.

۵. تحلیل تعارض میان محرمانگی امنیتی و حق شفافیت.

۶. تدوین سازوکار نظارت انسانی، ثبت تصمیم و امکان اعتراض.

نتیجه‌گیری

در جمع‌بندیِ این پژوهش، می‌توان گفت که به‌کارگیری هوش مصنوعی در حوزه امنیت ملی، اگرچه از منظر کارآمدی، سرعت تصمیم‌گیری، تحلیل کلان‌داده و پیش‌بینی تهدیدها ظرفیت‌های چشمگیری ایجاد می‌کند، اما هم‌زمان یک میدان پرتنش میان «امنیت» و «حقوق بنیادین» پدید می‌آورد. یافته‌های نظری نشان می‌دهد که مهم‌ترین پیامد این فناوری در سطح امنیت ملی، صرفاً افزایش توان عملیاتی دولت‌ها نیست، بلکه بازتعریف رابطه میان دولت، شهروند و داده است؛ رابطه‌ای که در آن مرزهای حریم خصوصی، شفافیت، نظارت‌پذیری و پاسخ‌گویی به‌شدت تحت تأثیر قرار می‌گیرد. از این‌رو، هوش مصنوعی را نمی‌توان صرفاً یک ابزار خنثی و فنی دانست؛ بلکه باید آن را یک فناوری حکمرانی‌ساز تلقی کرد که می‌تواند هم به ارتقای امنیت عمومی کمک کند و هم در صورت فقدان قواعد روشن، به نظارت فراگیر، تصمیم‌سازی غیرشفاف و تضعیف اعتماد عمومی بینجامد. بنابراین، نتیجه اصلی این است که امنیت ملیِ مبتنی بر هوش مصنوعی تنها زمانی مشروع و پایدار خواهد بود که در کنار کارایی، به اصول حقوقی و اخلاقی نیز وفادار بماند.

هوش مصنوعی در حوزه امنیت ملی کاربردهای گسترده و روبه‌گسترشی دارد و می‌تواند به‌عنوان ابزاری برای افزایش سرعت، دقت و دامنه تحلیل در اختیار نهادهای امنیتی قرار گیرد. یکی از مهم‌ترین این کاربردها، تشخیص و پیش‌بینی تهدیدهاست؛ به این معنا که سامانه‌های هوشمند با تحلیل الگوهای رفتاری، داده‌های تاریخی و نشانه‌های اولیه بحران، می‌توانند احتمال وقوع تهدیدهای امنیتی را زودتر از روش‌های سنتی شناسایی کنند. همچنین، تحلیل داده‌های امنیتی و اطلاعاتی در مقیاس بزرگ از دیگر کارکردهای مهم این فناوری است، زیرا هوش مصنوعی قادر است حجم عظیمی از داده‌های متنی، تصویری، صوتی و شبکه‌ای را در زمانی کوتاه پردازش کرده و الگوهای پنهان یا ارتباطات معنادار را استخراج کند. در حوزه دفاع سایبری نیز هوش مصنوعی نقش بسیار مؤثری دارد و می‌تواند در شناسایی حملات، کشف ناهنجاری‌ها، پاسخ خودکار به تهدیدات و تقویت سامانه‌های دفاعی به‌کار گرفته شود. افزون بر این، در پایش مرزها و شناسایی فعالیت‌های مشکوک، از ابزارهای هوشمند برای رصد رفت‌وآمدها، تشخیص رفتارهای غیرعادی و افزایش کارآمدی نظارت مرزی استفاده می‌شود. در نهایت، با گسترش عملیات روانی، جعل عمیق و انتشار اطلاعات نادرست، هوش مصنوعی به ابزاری مهم برای شناسایی، راستی‌آزمایی و مقابله با این تهدیدات نوین تبدیل شده است؛ هرچند همین ظرفیت‌ها اگر بدون نظارت و چارچوب حقوقی و اخلاقی به‌کار روند، می‌توانند خود منشأ چالش‌های تازه‌ای شوند.

در کنار ظرفیت‌های گسترده، به‌کارگیری هوش مصنوعی در امنیت ملی با تهدیدها و چالش‌های جدی نیز همراه است. مهم‌ترین چالش، نقض حریم خصوصی و گسترش نظارت فراگیر است؛ زیرا سامانه‌های هوشمند برای تحلیل تهدیدها معمولاً به گردآوری و پردازش حجم عظیمی از داده‌های شخصی و رفتاری نیاز دارند و این امر می‌تواند مرز میان امنیت

مشروع و کنترل ناموجه را مخدوش کند. چالش مهم دیگر، سوگیری الگوریتمی و تبعیض در تصمیم‌گیری است؛ به این معنا که اگر داده‌های آموزشی ناقص یا جهت‌دار باشند، خروجی سامانه ممکن است برخی گروه‌ها، مناطق یا رفتارها را ناعادلانه به‌عنوان تهدید تلقی کند و موجب تصمیم‌های نادرست یا تبعیض‌آمیز شود. افزون بر این، ابهام در مسئولیت حقوقی و اداری، یکی از مسائل اساسی است؛ زیرا در صورت بروز خطا یا خسارت، تعیین اینکه مسئول نهایی طراح، بهره‌بردار، ناظر یا نهاد تصمیم‌گیر است، دشوار می‌شود. از سوی دیگر، وابستگی بیش از حد به سامانه‌های خودکار می‌تواند به تضعیف نظارت انسانی، کاهش دقت قضاوت کارشناسان و حتی بروز خطاهای زنجیره‌ای در تصمیم‌گیری امنیتی بینجامد. همچنین، سوءاستفاده از هوش مصنوعی توسط بازیگران متخاصم برای تولید جعل عمیق، عملیات روانی، حملات سایبری پیشرفته و فریب افکار عمومی، تهدیدی نوظهور و بسیار مهم به شمار می‌رود. در مجموع، چالش اصلی این است که هوش مصنوعی در امنیت ملی اگر بدون چارچوب‌های حقوقی، اخلاقی و فنیِ دقیق به کار گرفته شود، می‌تواند به همان اندازه که ابزار تقویت امنیت است، به منبعی برای آسیب، بی‌اعتمادی و نقض حقوق بنیادین نیز تبدیل شود.

از منظر حقوقی، این پژوهش نشان می‌دهد که مهم‌ترین چالش، مسئله مسئولیت است؛ زیرا در سامانه‌های هوشمند، زنجیره تصمیم‌گیری به‌گونه‌ای توزیع می‌شود که تشخیص نقش دقیق طراح، تولیدکننده، بهره‌بردار، ناظر و نهاد سیاست‌گذار دشوار می‌گردد. در چنین وضعی، اگر یک سامانه امنیتی مبتنی بر هوش مصنوعی باعث خسارت شود، نمی‌توان به‌سادگی مسئولیت را به خودِ سامانه یا حتی فقط به یک عامل انسانی نسبت داد، بلکه باید سازوکاری روشن برای انتساب مسئولیت، جبران خسارت و نظارت قضایی وجود داشته باشد. همچنین، از آنجا که بسیاری از تصمیم‌های امنیتی در بستر محرمانگی اتخاذ می‌شوند، خطر کاهش شفافیت و محدود شدن حق اعتراض و دادرسی منصفانه افزایش می‌یابد. نتیجه این بخش آن است که نظام حقوقی باید پیشاپیش و نه پس از بروز بحران، برای این فناوری چارچوب‌های مشخصی مانند ارزیابی اثرات حقوق بشری، ثبت و مستندسازی تصمیم‌ها، نظارت انسانی مؤثر، و قواعد روشن مسئولیت مدنی و اداری تدوین کند. در غیر این صورت، هوش مصنوعی به جای آنکه ضامن امنیت باشد، خود می‌تواند به منشأ جدیدی از منازعه حقوقی و بی‌اعتمادی نهادی تبدیل شود.

از منظر اخلاقی و سیاست‌گذاری، نتیجه پژوهش بیانگر آن است که مسئله اصلی فقط «استفاده یا عدم استفاده» از هوش مصنوعی در امنیت ملی نیست، بلکه «چگونگی استفاده» از آن است. اگر این فناوری بدون معیارهای اخلاقی مانند تناسب، ضرورت، حداقل‌سازی آسیب، عدالت، عدم تبعیض و نظارت انسانی به‌کار گرفته شود، به‌تدریج کرامت انسانی را به حاشیه می‌راند و افراد را به داده‌هایی قابل پیش‌بینی و کنترل‌پذیر فرو می‌کاهد. در مقابل، اگر توسعه و استقرار آن بر پایه حکمرانی مسئولانه، ممیزی مستمر، قابلیت توضیح‌پذیری، و پاسخ‌گویی نهادی باشد، می‌تواند به ابزاری مشروع برای تقویت امنیت عمومی تبدیل شود، بدون آنکه به حقوق اساسی لطمه جدی بزند. بر همین اساس، نتیجه نهایی این پژوهش آن است که آینده مطلوب در این حوزه نه در رد کامل هوش مصنوعی، بلکه در «تنظیم‌گری هوشمند» آن نهفته است؛ تنظیم‌گری‌ای که میان امنیت ملی و آزادی‌های بنیادین توازن برقرار کند، از خطاهای الگوریتمی و سوگیری‌های ساختاری جلوگیری نماید، و اعتماد عمومی را به‌عنوان یکی از پایه‌های اصلی امنیت حفظ کند. به بیان دیگر، امنیت پایدار در عصر هوش مصنوعی، امنیتی است که با قانون، اخلاق و نظارت انسانی همراه باشد، نه امنیتی که صرفاً بر سرعت و قدرت فنی تکیه کند.

منابع:

Access Now. (2018). Human rights in the age of AI: A case study examining law enforcement use of AI-powered facial recognition. https://www.accessnow.org/wp-content/uploads/2018/11/AI-and-Human-Rights-Case-Study.pdf

Achuthan, K., & Ramanathan, S. (2024). Advancing cybersecurity and privacy with artificial intelligence: Current trends and future research directions. Nature Communications, 15, 11656524. https://doi.org/10.1038/s41467-024-53328-4

Achuthan, K., Ramanathan, S., Srinivas, S., & Raman, R. (2024). Advancing cybersecurity and privacy with artificial intelligence: Current trends and future research directions. Frontiers in Big Data, 7, 1497535. https://doi.org/10.3389/fdata.2024.1497535

Achuthan, K., Ramanathan, S., Srinivas, S., & Raman, R. (2024). Advancing cybersecurity and privacy with artificial intelligence: Current trends and future research directions. Frontiers in Big Data, 7, 1497535. https://doi.org/10.3389/fdata.2024.1497535

Agarwal, A., Sharma, R., & Singh, S. (2022). Artificial intelligence in border security: Applications, challenges and future directions. Defence Science Journal, 72(4), 513–522.

Ahmadi, A., Zargar, A., & Adami, A. (2023). Artificial intelligence technology and change in the national security of states. Political Defense Studies, 32(123), 39–64. https://journals.ihu.ac.ir/article_208054.html

Alipour, M., Nasiran Najafabadi, D., & Shirani, M. (2024). Civil liability arising from the use of artificial intelligence in the European Union. Fiqh Economic Studies Journal, 6(5), 12–20.

Allen, G. C., & Chan, T. (2017). Artificial intelligence and national security. Belfer Center for Science and International Affairs, Harvard Kennedy School.

Al-Sadat Maki, A., Al-Sadat Maki, Z., & Kashkoulian, E. (2024). A review of the liability resulting from artificial intelligence actions in the Iranian legal system. Scientific Journal of Fiqh, Law, and Criminal Sciences, 8(32), 71–79.

Babuta, A., Oswald, M., & Janjeva, A. (2020). Artificial intelligence and UK national security: Policy considerations. Royal United Services Institute. https://nrl.northumbria.ac.uk/id/eprint/42963/

Bagheri, P. (n.d.). Legal challenges of legal personality and civil liability of artificial intelligence. Journal of Private Law.

Bertolini, A. (2020). Artificial intelligence and civil liability. European Parliament, Policy Department for Citizens' Rights and Constitutional Affairs. https://www.europeansources.info/record/artificial-intelligence-and-civil-liability/

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., Hendrycks, D., Garfinkel, S., & others. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. arXiv. https://arxiv.org/abs/1802.07228

Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

Chen, Y., Zhang, J., & Li, X. (2023). Intelligent border surveillance using machine learning and multi-sensor fusion: A review. Sensors, 23(8), 3891.

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. California Law Review, 107(6), 1753–1820.

Council of Europe. (2024). Council of Europe framework convention on artificial intelligence and human rights, democracy and the rule of law. https://rm.coe.int/1680afae3c

Crawford, K. (2021). Atlas of AI: Power, politics, and the planetary costs of artificial intelligence. Yale University Press.

Crawford, K. (2021). Atlas of AI: Power, politics, and the planetary costs of artificial intelligence. Yale University Press.

Dai, D., & Boroomand, S. (2021). A review of artificial intelligence to enhance the security of big data systems: State-of-art, methodologies, applications, and challenges. Archives of Computational Methods in Engineering, 29, 1291–1309.

Dehajiesmaeili, M. (2024). Challenges of civil liability of artificial intelligence in the Iranian legal system: A look at regulations in the European Union. Government and Law Journal, 5(1), 81–98.

Dehghani Firoozabadi, S. J., & Chehrazad, S. (2023). Artificial intelligence and problematization of national security topics. Journal of Geopolitics, 19(2), 207–242. https://qpss.atu.ac.ir/article_16300.html

Dilmaghani, S. E., Brust, M. R., Danoy, G., Cassagnes, N., Pecero, J. E., & Bouvry, P. (2019). Privacy and security of big data in AI systems: A research and standards perspective. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 5737–5743).

ENISA. (2023). ENISA threat landscape 2023. European Union Agency for Cybersecurity. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Ghallab, A., & Elgohary, A. (2024). Artificial intelligence in national security: Opportunities and risks. Fari Journal of Global Affairs and Security Studies, 1(2), 101–121.

Gless, S., & Bertolini, A. (2022). Legal personality for artificial intelligence: A comparative perspective. Computer Law & Security Review, 44, 5–10.

Guo, Y., & Cai, Y. (2025). Model poisoning and prompt injection in large language models: Security risks and mitigation.

Hammoudeh, M., Newman, R., Denman, J., Ren, Q., & Khan, W. Z. (2020). Detection of suspicious activities using artificial intelligence in surveillance environments. IEEE Access, 8, 143222–143235.

Hekmatnia, M., Mohammadi, M., & Vateqi, M. (2019). Civil liability resulting from the production of autonomous artificial intelligence-based robots. Islamic Law Journal, 60, 249–273.

Hosseini, E., Al-Sadat, M., Amiri, M., & Khairallahi, M. A. (2024). Fiqh analysis of civil liability in artificial intelligence technology. Private Law Journal, 6(4), 13–29.

IBM Security. (2024). Cost of a data breach report 2024. IBM. https://www.ibm.com/reports/data-breach

Jerman-Blažič, B., & Klobućar, T. (2020). Removing the barriers in cross-border crime investigation by gathering e-evidence in an interconnected society. Information & Communications Technology Law, 29(1), 66–81. https://doi.org/10.1080/13600834.2020.1705035

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kalpande, P., Katre, P., & Madavi, C. (2025). A review of privacy and security in AI-driven big data systems: Standards, challenges, and future directions. International Journal of Research Publication and Reviews, 6(5), 1083–1096.

Kaur, R., Gabrijelčič, D., & Klobućar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, 101804. https://doi.org/10.1016/j.inffus.2023.101804

Klobućar, T., Kaur, R., & Gabrijelčič, D. (2024). Artificial intelligence for cybersecurity: Use cases and country perspective. In Lecture Notes in Networks and Systems (Vol. 1, pp. 117–126). Springer. https://doi.org/10.1007/978-3-031-54235-0_11

Lin, P., Abney, K., & Bekey, G. A. (2012). Autonomous military robotics: Risk, ethics, and design. In M. Anderson & S. Anderson (Eds.), Machine Ethics (pp. 38–53). Cambridge University Press.

Liu, Y., Wang, H., Zhang, J., & Chen, X. (2024). Deepfakes, misinformation, and national security: Emerging threats in the information environment.

Meunier, D., Ektefaie, Y., & Bouvry, P. (2021). AI-based border monitoring systems: Opportunities and limitations. Future Generation Computer Systems, 124, 112–123.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679. https://doi.org/10.1177/2053951716679679

National Security Agency, Cybersecurity and Infrastructure Security Agency, Federal Bureau of Investigation, Australian Signals Directorate, Australian Cyber Security Centre, et al. (2025). AI data security. https://media.defense.gov/2025/May/22/2003720601/-1/-1/0/CSI_AI_DATA_SECURITY.PDF

NIST. (2024). Artificial intelligence and cybersecurity: Current state and future directions. National Institute of Standards and Technology. https://www.nist.gov

Russell, S. J. (2019). Human compatible: Artificial intelligence and the problem of control. Viking.

Sangarsu, R. R. (2018). Navigating data security concerns with next-generation AI tools: A comprehensive review. International Journal of Scientific Research, 7(11).

Sarker, I. H., Kayes, A. S. M., & Badsha, S. (2021). Cybersecurity data science: An overview from machine learning perspective. Journal of Big Data, 8, 40. https://doi.org/10.1186/s40537-021-00472-7

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (pp. 59–68). ACM. https://doi.org/10.1145/3287560.3287598

Taddeo, M., McCutcheon, R., & Floridi, L. (2019). Trusting artificial intelligence in cybersecurity is a double-edged sword. Nature Machine Intelligence, 1, 566–567. https://doi.org/10.1038/s42256-019-0109-1

Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. Information Fusion, 64, 131–148. https://doi.org/10.1016/j.inffus.2020.06.014

U.S. Department of Defense. (2023, September 12). Contextualizing deepfake threats to organizations. https://media.defense.gov/2023/Sep/12/2003298925/-1/-1/0/CSI-DEEPFAKE-THREATS.PDF

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. Social Media + Society, 6(1).

West, J. D., et al. (2019). The role of artificial intelligence in combating misinformation. AI Magazine, 40(2), 88–97.

Yoganathan, V., [et al.]. (2025). AI-enabled cyber threats and defensive automation: Risks to critical infrastructure. [Journal details needed].

Zakerinia, H. (2023). The nature and basis of civil liability arising from artificial intelligence in Iranian and EU members' laws. Private Law, 20(1), 135–152. https://doi.org/10.22059/jolt.2023.356703.1007186

Zakerinia, H., & Gholampour, Z. (2024). Reasonable and conventional algorithms and strengthening the theory of attribution of civil liability to artificial intelligence. New Technologies Law Journal, 9, 156–168.

Zou, A., Wang, Z., Li, X., & [et al.]. (2024). Adversarial prompt injection attacks against large language models.