

## "Codice di guerra" di Mariarosaria Taddeo

di Alessandro Vecchio

29-12-2025

Recensione a: Mariarosaria Taddeo, *Codice di guerra. Etica dell'intelligenza artificiale nella difesa*, Raffaello Cortina Editore, Milano 2025, pp. 320, 25 euro (scheda libro)

Era il 2017 quando Putin dichiarò che chiunque sarebbe diventato leader nell'intelligenza artificiale avrebbe governato il mondo. Quello stesso anno usciva sul mercato l'assistente vocale Google Home e l'IA non aveva ancora il peso che ha oggi nel dibattito pubblico. Come spesso accade, tuttavia, l'innovazione tecnologica ha uno stretto rapporto con il mondo militare, e quindi nove mesi prima dell'uscita di ChatGPT la Russia invadeva l'Ucraina, anche facendo ricorso a droni e sistemi d'arma autonomi. Oggi questo non è più una sorpresa: sul fronte ucraino e lungo i confini dei Paesi della NATO i protagonisti sono i droni, e anche Israele si è servito di sistemi di riconoscimento di immagini per individuare bersagli a Gaza.

Ad aver studiato da subito l'espansione di questo fenomeno è Mariarosaria Taddeo, che negli ultimi anni ha lavorato assiduamente a temi che sono diventati per noi oggi cruciali: è lecito usare armi autonome? Su chi ricade la responsabilità in caso di errori? Taddeo ha raccolto oltre dieci anni di lavoro di ricerca in *Codice di guerra*, pubblicato in Italia ad agosto 2025 con il programmatico sottotitolo *Etica dell'intelligenza artificiale nella difesa*. E difatti il testo ruota intorno alla questione centrale dell'integrazione dell'intelligenza artificiale non solo nelle campagne militari, ma anche nelle fasi di progettazione, sviluppo, implementazione. Riferirsi alla difesa piuttosto che alla guerra non è per Taddeo una mera sostituzione retorica: i droni sono la punta dell'iceberg, ma, come viene esplicitato nella prefazione, lo spettro dei cambiamenti riguarda anche «il funzionamento delle organizzazioni della difesa, i processi decisionali e operativi, come i modi di acquisizione di dati e informazioni e le tattiche e strategie di guerra» (p. 13).

L'intera prefazione è una dichiarazione di intenti, a partire dall'obiettivo del libro: fornire un quadro di riferimento etico per contribuire alla governance dell'intelligenza artificiale nella difesa. Nel momento in cui in Europa si combatte una guerra con i sistemi d'arma autonomi (autonomous weapons systems, AWS) prima che sia concluso il dibattito etico sulla loro legittimità, diventa di massima urgenza un confronto accademico con le istituzioni e la società civile. E l'approccio di Taddeo è quello di un'etica translational, capace cioè di andare oltre ai livelli descrittivi e normativi, e di tradurli in linee guida per la progettazione e lo sviluppo. Così avviene nel corso del saggio, in cui analiticamente vengono considerati gli attuali scenari tecnologici e legislativi, seguiti da un ragionamento critico su ciò che va bene e su ciò che va cambiato. Il punto di vista dell'autrice guida la riflessione, in una conversazione continua con la letteratura esistente sul tema e uno sguardo puntuale sulle regolamentazioni esistenti. Taddeo si muove finemente in equilibrio tra i due estremi del suo lavoro, cioè l'assenza e l'eccesso di regole: se è vero che il primo è inaccettabile perché renderebbe possibile il caos, anche una regolamentazione eccessivamente puntigliosa che non tenga conto dell'attuazione pratica deve essere evitata.

Tra i due estremi oggi ci troviamo in quello inferiore, dunque l'assenza: non tanto di regole, ma di standard. Nel presentare le definizioni di AWS vengono mostrati dodici tentativi da parte di altrettante istituzioni, con esiti piuttosto differenti tra loro. Da qui il secondo obiettivo, che emerge più avanti nel corso del libro, ovvero quello di stimolare l'establishment della difesa al dibattito per l'individuazione e l'attuazione di principi etici che riguardano l'IA. Taddeo si confronta con la più immediata delle obiezioni: c'è bisogno di un'etica in guerra? È utile occuparsi dell'etica dell'IA in

questo campo? La risposta non solo è positiva, ma ribalta la questione giudicando indifendibili le due possibili posizioni di chi sale sul carro degli scettici: pacifisti e realisti. I primi sono coloro che giudicano la guerra un male assoluto, da evitare ad ogni costo, e ogni considerazione a riguardo pone il rischio di legittimarla; ma questa posizione non considera i casi di aggressione, come nella guerra in Ucraina, il cui diritto di difesa è innegabile. Il realismo invece non apprezza l'etica perché è vista come un ostacolo allo sviluppo e quindi come una debolezza nei confronti di un eventuale avversario che non se ne preoccupa. Ma, sostiene Taddeo, le democrazie liberali vivono in simbiosi con il diritto internazionale, e l'impiego dell'IA nella difesa con la contemporanea assenza di standard condivisi ha come conseguenza l'insanzionata elusione dei principi e dei valori fondanti delle stesse democrazie. L'etica dell'IA nella difesa non è dunque solo utile, ma fondamentale e necessaria per scoraggiare il ripresentarsi di atrocità come quelle delle guerre del Novecento.

Codice di guerra, come avverte la stessa Taddeo, non è un testo introduttivo: è necessaria una conoscenza tecnica almeno intermedia dell'intelligenza artificiale per sapersi districare in maniera autonoma tra i rischi, le problematiche e le opportunità a cui si fa riferimento; allo stesso modo è d'aiuto una conoscenza minima della "Teoria della Guerra Giusta", a cui si fa costante riferimento. Tuttavia, anche per chi ne fosse digiuno, la comprensione è facilitata dal fatto che il libro è, a tratti, divulgativo: probabilmente non l'intento iniziale dell'autrice, ma un effetto virtuoso che deriva da quello che è il pregio maggiore di questo saggio, ovvero la massima ambizione alla chiarezza. In ogni pagina si avverte l'impegno a fugare l'ambiguità e a correggere la vaghezza, in un tentativo encomiabile di mettere i puntini sulle i dove spesso sono proprio quest'ultime a mancare. È, in ultima istanza, un saggio che cerca di rendere chiaro ciò che chiaro non è, di mettere ordine dove appare solo un insieme di proposte confuse e dissonanti. Tutto ciò facilita la lettura e la comprensione, ma non si dimentichi che rimane la collezione di un lavoro accademico di alto livello svolto da Taddeo negli ultimi dieci anni all'Oxford Internet Institute; ciò significa un tono elevato nella discussione, un continuo rimando a definizioni, a regolamenti e a lavori di altri studiosi che si sono occupati negli anni dello stesso tema. La struttura argomentativa è quindi quella tipica di un paper accademico, ma l'accorgimento stilistico necessariamente realizzato per approdare nelle librerie ha raggiunto senza dubbio il suo scopo.

L'approccio analitico si rivela anche nella divisione in capitoli: i primi due sono introduttivi, e, in linea con la ricerca di chiarezza, fanno il punto su quelle che sono le caratteristiche costitutive dei sistemi di IA, evidenziando in che modo rappresentano rischi e opportunità quando applicate alla difesa. Come la capacità di calcolo è vitale per avere a disposizione informazioni rilevanti in momenti in cui sono necessarie decisioni rapide, la scarsa predicibilità dei sistemi IA - cioè la tendenza intrinseca a produrre risultati non attesi - rischia, se non controllata, di avere effetti devastanti. Molti esempi pratici rendono lampante l'urgenza della questione: se un sistema che cerca di individuare terroristi ha una percentuale di falsi positivi dello 0,008%, a prima vista appare come un risultato eccezionale, ma se applicato su una popolazione di 55 milioni di abitanti, risultano 4.400 persone ingiustamente etichettate dal sistema come potenziali terroristi. E quindi sorvegliate speciali, anche con le norme europee dell'AI Act che per le indagini sul terrorismo prevedono una deroga al divieto di sorveglianza. Immediata è l'associazione con gli AWS, in cui anche una soglia di errore molto bassa può significare uccisioni ingiustificate di civili. In questo senso, Taddeo insiste sulla necessità di un certo grado di controllo umano sempre garantito: l'essere umano deve avere la possibilità di intervenire in qualsiasi momento, ma soprattutto deve comprendere ciò che ha di fronte. Non è ammissibile una situazione nella quale un decisore militare si limiti a premere un pulsante delegando tutto il processo decisionale alla macchina. E qui si trova il punto di partenza: formazione IA a chiunque sia coinvolto nell'utilizzo di questi sistemi, dal generale al soldato semplice; allo stesso tempo, la riflessione etica deve partire dall'inizio, dalla fase di progettazione a quella di sviluppo e implementazione, e riguardare tutto il ciclo di vita dei sistemi IA: dalla fase 0 - l'idea - fino

---

all'audit etico che analizzi, una volta utilizzati, la loro funzionalità. Seguendo la linea dell'etica translational, Taddeo ragiona sulle modalità in cui ciò possa avvenire, proponendo linee guida e metodologie che siano riproducibili nella pratica. I principi etici non costituiscono leggi, né hanno valore normativo: devono quindi essere più simili a una bussola che a una mappa; devono orientare chi è coinvolto nel dominio di riferimento, ma ricordando che sono parte di un sistema e che la loro presenza acquista valore nel momento in cui diventa chiara e riproducibile la loro messa in pratica. In questo spostamento del focus etico dal che cosa al come, non bisogna però neanche perdere di vista i fondamenti teorici da cui scaturisce il dibattito, e su cui il testo ragiona a partire dalla "Teoria della Guerra Giusta", dottrina che da Sant' Agostino all'odierno diritto umanitario internazionale ha cercato di definire se una guerra può dirsi giusta e sotto quali condizioni. Quelli su cui fa riferimento Taddeo per tutto il corso del libro sono i principi di proporzionalità e di distinzione - tra combattenti e non combattenti. Entrambi sono messi in discussione dall'integrazione dell'IA, da un lato perché i sistemi attuali commettono ancora errori nel discernere i militari dai civili, dall'altro in quanto non c'è proporzione nel rischio se un militare umano nel campo di battaglia deve vedersela con un AWS, e dall'altro lato l'umano si limita alla supervisione a distanza. Il connotato cavalleresco di quest'ultimo principio è solo apparente; il presupposto è che chi accetta di entrare in guerra lo fa mettendosi in grande rischio, e si aspetta che nella fazione opposta accada lo stesso. Dimenticare questi due principi significa smettere di distinguere non solo i combattenti dai civili, ma anche lo stesso campo di battaglia, che sfuma fino a diventare indistinguibile dai territori non coinvolti. L'assunto di Taddeo è incisivo perché va oltre alla sacrosanta condanna da infliggere a uno Stato che attacca indiscriminatamente territori civili: ricorda che un'azione del genere non giustifica una reazione dello stesso tipo. Non secondo le regole del diritto internazionale: vale a dire, non secondo i valori delle democrazie liberali.

C'è dunque un'urgenza reale e significativa di standard etici: ma in quali usi? Quali sono i contesti di integrazione dell'IA nella difesa che più hanno bisogno di una riflessione a riguardo? Taddeo ne definisce tre, a livello di complessità crescente: gli usi come sostegno e supporto nella difesa, quindi le operazioni di intelligence, ad esempio, nella lotta al terrorismo; gli usi conflittuali e non cinetici, cioè la cosiddetta guerra informatica, volta alla sottrazione di informazioni sensibili attraverso attacchi hacker; se poi ci sono in gioco vite umane si parla di usi conflittuali e cinetici, come i droni e i sistemi d'arma autonomi. Ad avere più cose in comune, paradossalmente, sono il primo e il terzo, che condividono un impianto etico già impostato. Difatti l'IA applicata all'intelligence presenta grosso modo le stesse criticità dell'IA in altri campi: pregiudizi sistemici (bias), trasparenza nella raccolta dati, il sopracitato problema della predicibilità. A questo si aggiunge certamente il fatto che sia un apparato governativo ad usare i dati dei propri cittadini, che dovrebbero essere al sicuro: la cosiddetta intrusione è giustificata se i dati vengono solo raccolti e non esaminati? O non è mai ammissibile? E ancora: fino a che punto siamo disposti a lasciare autonomia decisionale all'IA? Si tratta certo di questioni spinose, ma che godono di ampia letteratura a riguardo e un già avviato dibattito. In parte è così anche per il terzo uso, quello più rischioso e su cui è necessario prestare maggiore attenzione, perché entrano in gioco vite umane. Gli usi conflittuali e cinetici racchiudono l'identificazione di bersagli nemici e l'utilizzo di AWS, quelle armi che godono di un certo grado di autonomia, come droni e carri armati, di cui Taddeo propone una definizione completa ed esaustiva. Qui la presenza di un impianto etico non riguarda la tecnologia ma la guerra: la "Teoria della Guerra Giusta" e il diritto internazionale sono punti di riferimento importanti, il cui aggiornamento è assolutamente necessario, ma rappresentano punti di partenza certamente validi. Non mancano però punti ciechi che Taddeo si impegna a illuminare, come la questione dell'assunzione di responsabilità quando viene usato un AWS. Rispondere che è dell'umano, per quanto un passo avanti, non è sufficiente: è del generale a capo, del ministro della difesa, di chi ha premuto il bottone? E come garantire che la presa di responsabilità porti ad un effettivo rispetto dei principi

---

etici? Taddeo si comporta da filosofa, provando innanzitutto a porre le giuste domande, quelle in grado di circoscrivere la questione, individuando i punti di interesse e scartando le vie senza uscita. Come già detto, tuttavia, il libro ha anche un'enorme spinta propositiva, e l'obiettivo dichiarato di fornire un quadro di riferimento etico nasconde quello implicito di volersi imporre come voce autorevole in un dibattito che in Europa ha massima centralità. Ciò è rivelato anche dalla trattazione svolta nei riguardi del secondo uso dell'IA nella difesa, quello conflittuale e non cinetico. In un mondo nel quale le informazioni, sotto forma di dati, acquisiscono piena centralità politica ed economica, le infrastrutture digitali diventano anch'esse campo di conflittualità, la cosiddetta cyberwarfare. Si pensi ad un attacco hacker nei confronti di un ospedale, nel quale i dati dei pazienti vengono manipolati: verrebbero somministrate medicine sbagliate finché non ci si rendesse conto dell'intrusione. Oggi un discorso analogo può essere fatto per diversi ambiti sensibili, tra cui il conto corrente e l'istruzione: è interesse di uno Stato in guerra indebolire il nemico dall'interno conquistando informazioni sensibili. Si potrebbe pensare che una guerra cibernetica sia preferibile rispetto ad una fisica, ma è un'obiezione miope che non tiene conto del fatto che la prima ha un ruolo subordinato e funzionale alla seconda, e che un'escalation sul piano digitale, raggiunto il suo limite, sfocia nel conflitto cinetico. Si tratta dunque di un tema tanto attuale quanto delicato, e la tesi di Taddeo è che non è possibile affrontarla tramite i principi etici correnti; significherebbe commettere lo stesso errore di chi, durante la Prima guerra mondiale, non si accorse che le nuove difese trincerate rendevano obsoleti gli attacchi tradizionali. In particolare, la deterrenza ha un ruolo molto particolare nella cyberwarfare, perché diventa difficile rispettare il principio di proporzionalità (non è così immediato stabilire l'entità di un danno informatico) e perché, al contrario della guerra cinetica, la tattica difensiva a lungo termine non può essere una strategia vincente. Tutto ciò richiede un cambio di paradigma, e neanche qui l'autrice si esime dal gravoso compito di tentarne un'impostazione, avventurandosi in un territorio dove per sua stessa ammissione la letteratura in merito è piuttosto scarsa.

È certo che l'uso più problematico e foriero di insidie è quello conflittuale e cinetico, che va regolato al più presto per evitare una degenerazione del fenomeno bellico, ed è infatti quello a cui Taddeo dedica più spazio, declinandolo in tre capitoli; è altrettanto vero però che la trattazione sulla cyberwarfare è ciò che rende questo libro unico: la guerra informatica viene messa sul piano che merita di avere, senza venir relegata a problema minore e secondario, né è discussa con inutili allarmismi. Anche qui entrano in gioco principi, definizioni e proposte, portate avanti con precisione e serietà. La serietà è uno dei tanti pregi di questo saggio, non solo nel tono, auspicabile in tale contesto; serio nel riconoscimento della gravità e dell'urgenza di una questione che non ha più bisogno di consigli e raccomandazioni, ma di una programmazione decisa e orientata all'azione. Taddeo, che svolge questo lavoro da molti anni con pieno riconoscimento del mondo accademico, pubblica adesso Codice di guerra per lanciare un monito alla società civile. È tempo di agire, ne va in gioco il concetto stesso di democrazia.