di Roberto Rosso

Qualsiasi tecnologia umana, qualsiasi artefatto è suscettibile degli usi più diversi, dai primi frammenti di roccia lavorati con cui si potevano costruire altri utensili, colpire la selvaggina o in alternativa colpire altri esseri umani. Ogni forma di comunicazione si presta, per definizione, a veicolare contenuti finalizzati a scopi diversi, a glorificare il potere nei suoi diversi aspetti, dalle stragi alle opere benefiche, la guerra, la sopraffazione o la solidarietà.
L'intelligenza umana da sempre si applica alla produzione delle armi, le stesse tecnologie, modelli organizzativi, strumenti di comunicazione si applicano alla cooperazione sociale o alla gestione delle guerre. L'apoteosi della capacità distruttiva è rappresentata dagli ordigni nucleari, mentre la produzione di energia elettrica tramite reattori a fissione si diffonde come alternativa all'uso delle fonti fossili, mentre proseguono le sperimentazioni per arrivare al controllo delle reazioni di fusione che conosciamo nella loro funzione distruttiva come bomba H. I trionfi della chimica a cavallo tra il XIX° e XX° secolo hanno prodotto l'uso dei gas nel primo conflitto mondiale, simboleggiate dall'Iprite -che prese il nome da Ypres[1]- il tioetere del cloroetano, detto anche gas mostarda, poi messi fuori legge. L'incubo dalla seconda metà del secolo XX° sono le armi biologiche, che sono state tirate in ballo nella diffusione, se non della creazione, del virus Sars-Cov2 che ha generato la più spaventosa pandemia dell'epoca moderna; le scienze della vita le biotecnologie sono alla base delle straordinarie possibilità di cura tanto quanto allo sviluppo delle armi biologiche.
L'avvento delle tecnologie digitali, di quell'ambio spettro che cadono sotto la definizione di Intelligenza Artificiale (I.A.) ha reso sempre più complesso, articolato e sfumato il confine tra l'uso legittimo, finalizzato al benessere, delle persone, delle comunità dell'umanità e quello finalizzato a colpire i diritti, l'integrità e la vita delle stesse. Del resto l'articolazione di questi confini, dove la distinzione tra il bianco e il nero sfuma n in mille sfumature di grigio, affonda nella natura dei rapporti sociali di produzione capitalistici, dominati a livello globale sotto diversi regimi politici.
L'attuale tendenza alla guerra, espressione di quel sistema, delle logiche di sfruttamento e competizione che lo fondano, le sue manifestazioni potenziali e attuali mostrano un utilizzo quanto mai diffuso ed avanzato di ogni filiera tecnologica, l'asservimento del sistema complessivo dell'innovazione tecnologica con un investimento di risorse mai visto prima.
La tendenza necrofila di questo sistema si esprime pienamente e compiutamente nella crisi climatica, che avanza senza che compaia all'orizzonte quel livello di solidarietà e cooperazione globali necessario a rallentare se non a fermare il processo di riscaldamento globale. Se questo è il contesto l'avvento dell'I.A. moltiplica le occasioni, le forme di uso duale delle tecnologie abilitando a questo la sua penetrazione pervasiva in ogni altra filiera tecnologica. Un esempio recente è il seguente.
Invertendo la finalità nell'utilizzo di un dispositivo di I.A. finalizzato alla scoperta di nuove molecole, in sole sei ore sono state proposte 40.000 nuovi composti tossici[2].
"In precedenza avevamo progettato un generatore commerciale di molecole de novo che abbiamo chiamato MegaSyn, che è guidato dalle previsioni del modello di apprendimento automatico della bioattività allo scopo di trovare nuovi inibitori terapeutici mirati a malattie umane. Questo modello generativo normalmente penalizza la tossicità prevista e premia l'attività target prevista. Abbiamo

semplicemente proposto di invertire questa logica utilizzando lo stesso approccio per progettare le molecole de novo, ma ora guidando il modello per premiare sia la tossicità che la bioattività.
(…)
Il software generativo sottostante è costruito su un software open source prontamente disponibile e simile ad esso. Per restringere l'universo delle molecole, abbiamo scelto di guidare il modello generativo verso composti come l'agente nervino VX, uno degli agenti di guerra chimica più tossici sviluppati nel corso del XX secolo.
(…)
Inoltre, non abbiamo sintetizzato fisicamente nessuna delle molecole; ma con una rete globale di centinaia di aziende commerciali che offrono sintesi chimiche, questo non è necessariamente un grande passo, inoltre questo settore è scarsamente regolamentato, con pochi o nessun controllo per prevenire la sintesi di nuovi agenti estremamente tossici che potrebbero potenzialmente essere usati come armi chimiche. È importante sottolineare che avevamo un essere umano coinvolto con una ferma voce morale ed etica del tipo "non sperimentate in quel campo" per intervenire. Ma cosa succederebbe se l'essere umano venisse rimosso o sostituito con un cattivo attore? Con le attuali scoperte e la ricerca sulla sintesi autonoma, un ciclo completo di progettazione-produzione-test applicabile alla produzione non solo di farmaci, ma anche di tossine, è a portata di mano.
(…)
Il software di machine learning open source è la strada principale per l'apprendimento e la creazione di nuovi modelli come il nostro, mentre i dataset sulla tossicità -che forniscono un modello di base per le previsioni per una serie di obiettivi relativi alla salute umana- sono prontamente disponibili."

Degli stessi autori, prendendo in esame il medesimo esperimento, è l'apello'AI in drug discovery: A wake-up call'[3] nel quale l'accento è posto sulle potenzialità opposte insite nell'uso dell'I.A., ma propone una riflessione sul 'lato oscuro' della ricerca in tutti i campi.
"Il nostro esperimento proof-of-concept solleva sfide per l'intera comunità che utilizza l'IA generativa per progettare molecole. Come utilizziamo le nostre conoscenze per gestire il duplice uso dell'IA La scoperta di nuovi farmaci è importante. Una reazione istintiva sarebbe quella di bloccare i dati e modelli, o per lo meno di imporre controlli regolamentari. Scienziati del Progetto Manhattan si resero conto che "non potevano rimanere in disparte di fronte alle conseguenze del loro lavoro" e distruzione di Hiroshima e Nagasaki, cominciarono a impegnarsi in una serie di iniziative scientifiche responsabili. Mentre storicamente, le iniziative scientifiche responsabili tendono ad essere adottate dopo una tragedia, il campo della scoperta di farmaci ha un'opportunità unica mettere in atto misure preventive prima che un esempio reale di uso improprio porti a catastrofe. Ora è il momento, per noi e per le molte centinaia di gruppi e aziende che utilizzano approcci di IA per la scoperta di farmaci, per trovare il nostro modo di scienza in modo responsabile e a proporre iniziative in grado di ridurre il rischio di riprodurre di esiti dannosi."
L'esperimento che ha evidenziato i rischi impliciti nell'uso dell'I.A. nel settore farmaceutico è stato condotto dal gruppo di ricercatori a partire da un invito dell'Istituto federale Svizzero per la protezione NBC (Nucleare, biologico, Chimico) che li ha condotti ad utilizzare il dispositivo di machine learning MegaSyn, creato per formulare nuovi farmaci, allo scopo inverso creare formule di

composti tossici, sino ad allora sconosciuti molti dei quali neppure presenti nei repertori dei composti messi al bando.
Se confrontiamo il processo messo in atto in questo caso- vale a dire l'intenzionalità e la facilità con cui si è giunti alla formulazione di decine di migliaia di composti tossici e nocivi- ci si rende conto dell'inadeguatezza di qualsiasi legge, direttiva o regolamento mirante a classificare i livelli di rischio presenti nell'utilizzo dell'I.A. ed in base a questa classificare tenerne sotto controllo la realizzazione.
Il riferimento è all'I.A. Act europeo che entrerà in vigore nel 2026, che intende offrire un quadro di riferimento quando più esaustivo possibile, ma si deve confrontare con l'evoluzione esponenziale delle tecnologie, la proliferazione quindi dei campi di utilizzo che possono rispondere ai più diversi interessi. Nel quadro della competizione economico-finanziaria e geopolitica, come nel caso del cambiamento climatico, siamo in assenza di una cooperazione globale tra i protagonisti di questo straordinario balzo in avanti tecnologico, benché ognuno proponga il proprio modello regolamentare l'IA. Act europeo, l'Ordine Esecutivo del presidente Biden -che in quanto tale potrebbe essere ritirato e contraddetto da un altro presidente- la strategia di controllo del governo cinese nei confronti dei protagonisti del mercato digitale, subordinato quindi agli obiettivi delle istituzioni statali.
In Europa la redazione dell'I.A. Act è stato preceduto da una serie si studi, tra cui il documento intitolato 'ORIENTAMENTI ETICI PER UN' IA AFFIDABILE'[4], prodotto dal Gruppo di esperti ad alto livello sull'intelligenza artificiale (HLEG AI), nel quale sono contenuti tutti gli interrogativi ed i riferimenti che hanno portato poi all'emissione della direttiva. Come abbiamo scritto[5] i risultati dei lavori dell'HLEG AI sono stati presentati alla prima Assemblea europea dell'IA nel giugno 2019. A seguito dell'Assemblea, la Commissione europea ha prorogato di un altro anno il mandato del gruppo. Questo mandato esteso ha permesso al gruppo di aumentare il proprio lavoro e portare a termine le Linee guida etiche per una IA affidabile. Il mandato del gruppo ad alto livello di IA si è concluso nel luglio 2020 con la presentazione di altri due elementi L'elenco di valutazione finale per l'IA affidabile (ALTAI)[6] Considerazioni settoriali sulle raccomandazioni politiche e di investimento.
Indubbiamente si è arrivati all'elaborazione dell'A.I. Act con un importante lavoro preparatorio, tuttavia nel frattempo il contesto tecnologico, finanziario, energetico ed applicativo è stato letteralmente rivoluzionato.
L'ALTAI specifica i 7 requisiti già individuati
1 Intervento e sorveglianza umani
Inclusi i diritti fondamentali, l'intervento umano e la sorveglianza umana
2 Robustezza tecnica e sicurezza
Inclusi la resilienza agli attacchi e la sicurezza, il piano di emergenza e la sicurezza generale, la precisione, l'affidabilità e la riproducibilità.
3 Riservatezza e governance dei dati
Inclusi il rispetto della riservatezza, la qualità e l'integrità dei dati e l'accesso ai dati.
4 Trasparenza
Incluse la tracciabilità, la comprensibilità e la comunicazione
5 Diversità, non discriminazione ed equità
Incluse la prevenzione di distorsioni inique, l'accessibilità e la progettazione universale, e la partecipazione dei portatori di interessi

6 Benessere sociale e ambientale
Inclusi la sostenibilità e il rispetto ambientale, l'impatto sociale, la società e la democrazia
7 Accountability
Inclusi la verificabilità, la riduzione al minimo degli effetti negativi e la loro segnalazione, i compromessi e i ricorsi.
Il punto cruciale della discussione successiva nelle istituzioni europee per arrivare all'approvazione dell'A.I. Act è stato ovviamente il nesso tra controllo e possibilità, libertà di sviluppo della tecnologia, in un quadro di competizione forsennata e di sviluppo esponenziale della tecnologia stessa.
L'aspetto paradossale dell'esempio portato nella creazione di nuovi farmaci o composti tossici è il fatto che esso si sia basato sulla disponibilità di database aperti dei composti chimici e sull'uso di software commerciali o addirittura liberi; siamo ben lontani dell'enorme investimento di risorse in termini di dati e risorse computazionali utilizzate per addestrare i Large Language Modules di ultima generazione. La fallibilità, la debolezza delle procedure di controllo quindi deriva sia dall'enormità delle risorse impiegate nei LLM -siamo in attesa dell'uscita per l'estate di ChatGPT-5 con i suoi miliardi di parametri di controllo aggiuntivi rispetto al modello precedente- tanto quanto dalla facilità con cui con minori mezzi si possono ottenere risultati importanti tanto positivi quanto negativi.
Nel documento Dual-Use AI Technology in China, the US and the EU – *Strategic Implications for the Balance of Power*[7] prodotto da un gruppo di ricercatori del The Peace Research Institute Oslo (PRIO) si afferma.
"A differenza delle trasformazioni militari del passato, in cui l'innovazione spesso emergeva dall'interno militare-industriale, oggi il settore IT civile è all'avanguardia nei cicli di innovazione. Le forze armate di tutto il mondo stanno facendo a gara per stare al passo con i cicli dell'innovazione. Ospitano la costante preoccupazione che le tecnologie dirompenti rendano obsolete le strategie consolidate nel tempo, da qui la corsa per applicare e integrare le tecnologie di intelligenza artificiale all'interno della pianificazione militare a un ritmo più veloce rispetto ai propri concorrenti e avversari.
Perché i motori cruciali di questo processo sono le tecnologie civili – hardware e software – la questione di come mantenere il controllo sui prodotti commerciali, in modo che non possano essere utilizzati in campo militare dagli avversari, è in cima all'agenda politica dei funzionari governativi che hanno operano nel campo della difesa e degli affari esteri. Il duplice uso delle tecnologie, a lungo un pilastro delle discussioni sulla tecnologia nucleare, ha un significato e un'urgenza completamente nuovi nel contesto della tecnologia militare potenziata dall'IA. Considerando che con il nucleare le armi sono difficili da costruire e trasportare senza essere scoperte, diverso è il caso delle nuove tecnologie digitali costituite da linee di codice, da algoritmi che possono essere facilmente trasferite da un punto all'altro. D'altra parte l'uso di queste tecnologie si basa su infrastrutture fisiche, da qui la crescente pressione per restringere la circolazione di tecnologie di rilevanza strategica come computer avanzati, dispositivi per la creazione dei chip e così via. Poiché queste tecnologie sono pervasivamente implicate ed utilizzate nelle comunicazioni civili nei settori dell'IT e sono fondamentali per il libero flusso di beni e servizi in tutto il mondo, le restrizioni sull'ampia gamma di prodotti che possono essere classificati come "a duplice uso" sollevano questioni impegnative che la presente relazione si prefigge di affrontare."

In queste righe si evidenzia l'inversione dei ruoli tra il settore civile e quello militare, il che implica però, nell'enorme sviluppo globale delle spese militari, uno straordinario impegno dei diversi stati ed alleanze nel trasferire nel campo militare le tecnologie sviluppate in campo civile.
Se l'analisi condotta in questo documento non è riassumibile, è del tutto evidente l'interrelazione che si stabilisce nell'uso delle tecnologie digitali nella competizione globale nel campo economico ed in quello militare, laddove i due regimi competitivi e le rispettive aree di influenza nelle diverse regioni del globo sono strettamente correlati.
"In generale, è facile vedere come la tecnologia plasmi l'equilibrio di potere sia attraverso mezzi economia che militari, ad esempio incidendo sul potere economico di un paese e influenzando un la capacità del paese di condurre la guerra; tuttavia, è meno chiaro in che modo specifiche applicazioni di IA possano tradursi nel potere militare. Ciononostante, possiamo fare alcune ipotesi sulle sfide che l'IA pone per le stabilità. Horowitz sostiene che le applicazioni dell'IA hanno il potenziale per plasmare i conflitti futuri in un una serie di macro aree, ad esempio aumentando la velocità con cui i paesi possono combattere, e se da un lato vi è incertezza su specifiche applicazioni militari dell'IA, dall'altro un'accelerazione del ritmo nella conduzione della guerra può sconvolgere in modo significativo le strutture organizzative. I sistemi di IA potrebbero portare anche a cambiamenti nella strategia militare, ad esempio sostituendo le macchine agli esseri umani nel prendere determinate decisioni. Inoltre, la percezione di qualsiasi progresso nel campo della dual-use technology rischia di creare un " dual-use security dilemma ". Questo può succedere anche agli attori che mirano a sviluppare la tecnologia solo per scopi civili e che possono trovarsi in una situazione di corsa agli armamenti a causa del fatto che altri attori potrebbero considerare tale tecnologia come una minaccia per le loro sicurezza."
Questo è un altro assunto fondamentale dell'analisi che viene ulteriormente approfondito ed esemplificato, vale a dire il processo di radicale e veloce trasformazione tecnologica di ogni rapporto sociale e filiera produttiva indotta dall'I.A*. si traduce in un regime di instabilità dei rapporti di forza degli stessi terreni di confronto in campo militare, rendendo quindi instabili i rapporti geopolitici, il confronto geostrategico;* nel campo militare si coniugano la precisione nell'individuazione e nell'annientamento di un singolo obiettivo, considerato strategico, con l'annientamento di massa delle strutture militari e delle truppe nemiche assieme alla distruzione delle infrastrutture che reggono la compagine sociale nemico e le campagne di terrore nei confronti delle popolazioni coinvolte nel conflitto. Di tutto questo le cronache offrono tragici esempi giorno dopo giorno.
Va da sé che la capacità dei nuovi software fondati sulla tecnologia transformer di produrre immagini, video e suoni che imitano perfettamente la realtà aggiunge uno straordinario campo di incertezza in ogni ambito della comunicazione, delle relazioni sociali, uno straordinario strumento nel campo della cosiddetta guerra ibrida che interviene nel campo delle relazioni interne, della comunicazione, della condivisione delle conoscenze delle formazioni sociali avversarie, falsificandone i contenuti.
L'uso duale della tecnologia, come si diceva nell'introduzione, propone diversi livelli di pericolosità che toccano il loro acme nel campo militare, ma esiste una gradazione di utilizzi che attraversa tutte le forme di competizione, di esasperazione delle diseguaglianze e delle forme di sfruttamento in società che sulla competizione le diseguaglianze si fondano.
La letteratura disponibile sul tema del 'dual use technology' nel campo dell'I.A. è sterminata, quanto

citato è solo un piccolissimo assaggio, che pure illumina i termini della questione nel loro rapido evolversi. Queste brevi note intendono contribuire all'approfondimento del tema, all'apertura di un confronto pubblico che sino ad ora si è svolto in modo sotterraneo e riguarda innanzitutto il mondo della ricerca in tutte le sue branche e articolazioni, dalle imprese, ai centri di ricerca, alle università e coinvolge tutti i campi disciplinari nessuno escluso poiché si parla della manipolazione della vita, delle relazioni umane, dei rapporti sociali in ogni loro aspetto resa possibile in modi sino ad oggi impensati dalle tecnologie dell'I.A., un confronto che va ben oltre la questione dei rapporti con le università ed i centri di ricerca israeliani, oggi all'ordine del giorno e riguarda tutto intero il futuro delle nostre società, la possibilità stessa di sopravvivenza dell'umanità.

Roberto Rosso

1. https://it.wikipedia.org/wiki/Seconda_battaglia_di_Ypres https://it.wikipedia.org/wiki/Iprite.[↩]
2. https://www.nature.com/articles/s42256-022-00465-9.epdf?sharing_token=GFRdRARfbikix_awLoa_w9RgN0jAjWel9jnR3ZoTv0M6VuGuVWKcBJFL5U5ocXOA5zcnGmZOUPQzouuai7vI0VssBKnaLbKdoBb2D8bZtqxuf8yx6_vtqooz7wBUrUztawrbCha4sM-QzvjBd9eutMKp3omle4YxqXFL_SOXG-TyfgN2DDkzSTlX8GBQ9-xTETGI0mQVvM412mwFQLnqS-aM3Ta43CFWZgfo78jMTKIU3_OUkAImu7Srs9RCbLF_hS57ISWsnEFb1FhB9aIKKw%3D%3D.[↩]
3. https://www.sciencedirect.com/science/article/abs/pii/S1359644622004032.[↩]
4. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.[↩]
5. https://transform-italia.it/regolare-lintelligenza-artificiale-o-liberare-lintelligenza-sociale/.[↩]
6. https://digital-strategy.ec.europa.eu/it/node/806.[↩]
7. https://www.prio.org/publications/13150.[↩]