

*I quaderni di*  
Agenda  **Digitale** <sup>eu</sup>

GENNAIO-APRILE 2024

n. 0016

Agendadigitale.eu è una testata scientifica e giornalistica registrata al Tribunale di Milano  
Dati di riferimento

Iscrizione ROC n. 16446

ISSN 2421-4167

Numero registrazione 1927, Tribunale di Milano

Editore: Digital360

Focus e ambito:

La rivista scientifica, i Quaderni di Agendadigitale.eu, pubblica fascicoli quadrimestrali in open access.

Lo scopo è creare un luogo per accompagnare i passi dell'Italia verso la necessaria rivoluzione digitale, con approfondimenti multidisciplinari a firma di esperti delle materie afferenti all'Agenda Digitale italiana ed europea

Submission e norme editoriali

Per effettuare una submission è necessario concordare prima un argomento e le misure precise contattando [info@agendadigitale.eu](mailto:info@agendadigitale.eu).

Inviare un abstract di circa 500 caratteri alla testata, presentando l'articolo.

Le misure del testo finale saranno comprese tra 6mila e 20mila caratteri, salvo accordi per misure superiori.

I riferimenti bibliografici dovranno essere preparati in conformità alle regole dell'APA style, 6a edizione (si vedano le linee guida e il tutorial).

Gli autori sono invitati a tener conto degli articoli già pubblicati nella rivista e di citarli nel loro contributo qualora siano ritenuti di interesse per il tema trattato.

**Comitato scientifico****Presidente:**

Alessandro Perego, Politecnico di Milano

**Membri del Comitato scientifico**

Francesco Agrusti, Università degli Studi Roma TRE

Davide Bennato, Università di Catania

Giovanni Biondi, Indire, Iulm

Giovanni Boccia Artieri, Università di Urbino

Paolo Calabrò, Università Vanvitelli di Caserta

Antonio Chella, Università di Palermo

Stefano Cristante, Università del Salento

Lelio Demichelis, Università Insubria

Marco del Mastro, Unicusano

Carlo Alberto Carnevale Maffè, Università Bocconi di Milano

Carmelo Cennamo, Università Bocconi di Milano

Michele Colajanni, Università degli Studi di Modena e Reggio Emilia

Mariano Corso, Politecnico di Milano

Ottavio Di Cillo, università di Bari

Maurizio Ferraris, università di Torino

Ivan Ferrero, psicologo

Paolo Ferri, Università Bicocca di Milano

Pietro Fiore, Università di Foggia

Stefania Fragapane, Università degli Studi di Enna Kore

Alfonso Fuggetta, Politecnico di Milano

Alberto Gambino, Università Europea di Roma

Carlo Giovannella, Università Tor Vergata di Roma

Renato Grimaldi, Università di Torino

Mariella Guercio, Università Sapienza di Roma

Mauro Lombardi, Università di Firenze

Mariano Longo, Università del Salento

Roberto Maragliano, Università Roma Tre

Massimo Marchiori, Università di Padova

Berta Martini, Università di Urbino Carlo Bo

Leonardo Menegola, università Milano Bicocca

Tommaso Minerva, Università degli studi di Modena e Reggio Emilia

Mario Morcellini, Università degli Studi di Roma "La Sapienza"

Giuliano Noci, Politecnico di Milano

Fabrizio Onida, Università Bocconi di Milano

Norberto Patrignani, Politecnico di Torino

Mario Pireddu, Università degli Studi della Tuscia

Franco Pizzetti, Università di Torino  
Alessio Plebe, Università di Messina  
Roberto Pozzetti, psicanalista, LUDeS Campus Lugano, università Insubria  
Antonio Rafele, Università di Parigi (CEAQ- Université Paris Descartes La Sorbonne)  
Francesco Sacco, Università Bocconi di Milano  
Donatella Sciuto, Politecnico di Milano  
Nicola Strizzolo, Università di Udine  
Elena Valentini, Università Sapienza di Roma  
Guido Vetere, Università Sapienza di Roma

### **Comitato di referaggio**

Coordinatore: Luca Gastaldi, Polimi

Mauro Andreolini, sicurezza informatica, Unimore

Luca Baccaro, concorrenza, diritto comunicazioni elettroniche e dei media; studio legale Lipani  
Catricalà & Partner

Raffaello Balocco, IT e innovazione, Politecnico di Milano

Francesco Capparelli, privacy, cyber security, ecommerce, data management, identità digitale;  
studio legale ICT Legal Consulting

Antonio Chella, ingegneria informatica, intelligenza artificiale, Università di Palermo

Marco Centorrino, Università di Messina – processi culturali e comunicativi, nuove tecnologie

Ida Cortoni, media education e digital literacy; Dipartimento di Comunicazione e Ricerca Sociale,  
Sapienza Università di Roma

Giuseppe D'Acquisto, Autorità garante privacy, sicurezza e privacy

Mario dal Co, Economista e manager, già direttore dell'Agenzia per l'innovazione

Lelio Demichelis, Università Insubria, sociologia, economia

Francesco Di Giorgi, diritto dell'informazione e della comunicazione, tutela dei consumatori, diritto  
delle comunicazioni elettroniche; Agcom

Leonella Di Mauro, data management, e-commerce, tutela del consumatore, diritto delle  
comunicazioni elettroniche; Agcom

Luisa Franchina, cyber security, Hermes Bay

Luca Gastaldi: eGov, sanità, telecomunicazioni, procurement pubblico, design thinking, Smart  
Working, Politecnico di Milano

Maurizio Gentile, professore associato, Università di Roma LUMSA, didattica e pedagogia

Antonio Ghezzi: strategia, business model, startups, mobile, Politecnico di Milano

Ugo Imbriglia, sociologo

Gevisa La Rocca, Università Kore di Enna, piattaforme digitali, communication research, analisi  
qualitativa dei dati

Nicola La Sala, registro degli operatori della comunicazione, fattura elettronica, industria4.0,  
editoria, cittadinanza digitale; Agcom

Emanuele Lettieri, sanità Politecnico di Milano

Maria Beatrice Ligorio, psicologia, università di Bari

Marika Macchi, economia, Unifi

Riccardo Mangiaracina: fatturazione elettronica, eCommerce, logistica e trasporti, export, Politecnico di Milano

Mirco Marchetti, Sicurezza informatica, unimore

Chiara Marzocchi, economia, Università di Manchester

Cristina Masella, Sanità, Politecnico di Milano

Carmelina Maurizio, Dipartimento di Filosofia e Scienze dell' Educazione Università di Torino

Stefano Moriggi, scienze della comunicazione, filosofia, Bicocca di Milano

Davide Mula, sanità digitale, cyber security, privacy; Agcom

Simone Mulargia, internet and social media studies; Lumsa

Antonella Napoli, sociologia, media e comunicazione, giornalista

Sebastiano Nucera, Università di Messina, Media e Tecnologie Indossabili

Achille Pierre Paliotta, Social cybersecurity, disinformazione, tecnologie digitali, intelligenza artificiale, sociologia economica; INAPP

Francesco Paoletti, docente di organizzazione aziendale e gestione delle risorse umane, Università degli Studi di Milano-Bicocca

Norberto Patrignani, computer ethics, filosofia, Politecnico di Torino

Dunia Pepe, Inapp e Università Roma Tre, cultura e formazione digitale

Alessio Plebe, Università di Messina, Scienze cognitive, pedagogiche, psicologiche

Francesco Pira, Unime, comunicazione pubblica, le dinamiche social, le fake news e i processi di disinformazione

Franco Pizzetti, diritto, privacy, università di Torino

Barbara Quacquarelli, scienze umane e formazione, università Milano Bicocca

Antonio Rafele, Sociologia dei processi culturali e comunicativi, Unicusano

Filippo Renga: turismo digitale, smart agrifood, finance and banking, mobile, Politecnico di Milano

Angelo Rovatti, tutela del diritto d'autore, diritti connessi, Diritto dei media; Agcom

Christian Ruggiero, sociologia del giornalismo e comunicazione politica; Dipartimento di Comunicazione e Ricerca Sociale, Sapienza Università di Roma

Franco Torcellan, Associazione RED – Laboratorio di Ricerca Educativa e Didattica “Formare Trasformare Innovare”

Angela Tumino: Internet of Things, logistica e trasporti, smart city, Politecnico di Milano

Simone Vannuccini, economia, SPRU

Francesco Varanini, filosofia, formazione, università di Pisa

Guido Vetere, Università Sapienza di Roma, intelligenza artificiale, tecnologia

## INDICE DEL FASCICOLO

<b>Alla ricerca della coscienza artificiale: sfide e nodi etici .....</b>	<b>7</b>
Di <b>Riccardo Manzotti</b> , Ordinario di Filosofia Teoretica, IULM, Milano	
<b>Autopsia psicologica dell'autore di reato: un modello informatico per decifrare il movente .....</b>	<b>14</b>
Di <b>Vincenzo Lusa</b> , Fellow American Academy of Forensic Sciences-Università Sapienza (Roma)	
<b>Prove tecniche di AI a scuola: come integrarla nel nostro sistema educativo.....</b>	<b>24</b>
Di <b>Daniela Di Donato</b> , Docente di italiano (Liceo scientifico), PhD in Psicologia sociale, dello sviluppo e della Ricerca educativa presso Sapienza Università di Roma, esperta di metodologie didattiche, inclusione e uso delle tecnologie digitali a scuola	
<b>Terrorismo islamico: spazialità reali e virtuali .....</b>	<b>31</b>
Di Marino D'Amore, Docente di Sociologia generale presso Università degli Studi Niccolò Cusano	
<b>Guerre di algoritmi: sistemi d'arma autonomi e agentività digitale.....</b>	<b>42</b>
Di <b>Giuseppe Galetta</b> , funzionario presso l'Università degli Studi di Napoli Federico II	
<b>L'effetto degli smartphone sulle funzioni esecutive di bambini e adolescenti: cosa dicono gli studi .....</b>	<b>62</b>
Di <b>Giulia Conte</b> , Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma, <b>Dario Esposito</b> , Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma, <b>Mauro Ferrara</b> , Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma, <b>Sara Romano</b> , Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma	
<b>Ri-organizzare la Realtà Estesa: il ruolo flessibile della cultura e delle simulazioni filosofiche e tecnologiche .....</b>	<b>72</b>
Di <b>Carmine T. Recchiuto</b> , Università degli Studi di Genova, DIBRIS, RICE Lab, <b>Lorenza Saettone</b> , Filosofa specializzata in Epistemologia e Cognitivism, PhD Student in Robotics and Intelligent Machines for Healthcare and Wellness of Persons e <b>Antonio Sgorbissa</b> , Università degli Studi di Genova, DIBRIS, RICE Lab	
<b>Perché l'AI non potrà mai replicare l'essenza dell'intelligenza umana .....</b>	<b>87</b>
Di <b>Roberto Pozzetti</b> , Psicoanalista, Professore a contratto LUDeS Campus Lugano, Professore a contratto Università dell'Insubria, autore del libro 'Bucare lo schermo. Psicoanalisi e oggetti digitali', già referente per la provincia di Como dell'Ordine degli Psicologi della Lombardia	
<b>L'AI può aiutarci a "leggere" i pensieri? Progressi concreti e dubbi etici .....</b>	<b>91</b>
Di <b>Tommaso Ciulli</b> , Psicologo	

# Alla ricerca della coscienza artificiale: sfide e nodi etici

L'evoluzione dei modelli di linguaggio (LLM) basati su algoritmi Transformer ha intensificato il dibattito sulla coscienza artificiale. Studi e teorie si dividono sull'interpretazione della coscienza e sulla sua possibile emulazione nell'IA. Nonostante il crescente interesse e le ricerche, la questione rimane aperta, tra approcci scientifici divergenti e profonde implicazioni etiche

Di Riccardo Manzotti, Ordinario di Filosofia Teoretica, IULM, Milano

A partire dal 2017, l'irrompere dei modelli di linguaggio di grandi dimensioni o LLM basati su algoritmo Transformer (Vaswani, Shazeer et al. 2017) ha posto con sempre maggiore urgenza il problema della coscienza. Riviste autorevoli come *Nature* hanno dato voce alla richiesta, avanzata da molti, di rispondere alla domanda se questo tipo di IA abbia o meno qualcosa di simile alla coscienza (Lenharo 2023). Il fatto che l'IA disponga di una competenza linguistica che, fino a ieri, era considerata la manifestazione di una mente cosciente, pone la domanda: l'IA è o sta per diventare cosciente?

<https://www.agendadigitale.eu/industry-4-0/intelligenza-artificiale-non-avra-mai-coscienza-ma-potra-cambiare-la-nostra-ecco-come/>

## Il dibattito sulla coscienza dell'intelligenza artificiale

La domanda è, non da oggi, al centro delle ricerche e delle domande di numerosi centri di ricerca. In Italia, Antonio Chella, professore di robotica presso il dipartimento di informatica dell'università di Palermo, è impegnato su questo tema ed è anche l'editor della rivista *Journal of AI and Consciousness*. In tempi ormai lontani, insieme con Vincenzo Tagliascio ci si interrogava sulla possibilità di costruire una macchina cosciente (Manzotti e Tagliascio 2001) e l'interrogativo è stato presente fin dall'inizio degli studi sulla intelligenza artificiale (Turing 1950). Oggi, autori prestigiosi come Noam Chomsky, Lenore Blum, David Chalmers stanno affrontando lo stesso problema (Chalmers 2022, Chomsky, Roberts et al. 2023, Melloni, Mudrik et al. 2023). Classici approcci all'intelligenza naturale come il paradigma dell'*inferenza attiva* di Karl Friston vengono oggi confrontato con le capacità degli LLM (Pezzulo, Parr et al. 2024). Ovviamente, la ricerca sulla coscienza artificiale implica, a monte, una risposta circa la natura della coscienza. Non possiamo trarre conclusioni sulla possibilità che una macchina sia cosciente, senza avere un'idea, anche molto approssimativa, di come gli esseri umani (e molti animali) siano capaci di fare esperienza. Per questo motivo, per parlare di coscienza artificiale si deve partire dal problema più ampio: la natura e localizzazione della coscienza nel mondo fisico.

## Perché è così difficile definire la coscienza

Proviamo a definire il problema della coscienza e perché sia così difficile (anzi è definito da trent'anni *il problema difficile o hard problem*). Sarò breve. La coscienza è generalmente intesa come la capacità di fare esperienza di quello che succede. Un sasso non prova dolore, un essere umano sì. Una telecamera non vede i colori (anche se li registra e memorizza), un essere umano sì. In che cosa consiste questa capacità? La risposta breve è che nessuno oggi può rispondere in termini oggettivi e fisici anche se tutti ne intuimo il significato nel momento in cui ci sottoponiamo a un intervento chirurgico. In quel momento prima che il bisturi incida la nostra pelle, l'eliminazione della coscienza grazie all'anestesia non è più qualcosa di vago. È qualcosa di reale, molto reale. Tuttavia, si tratta di una realtà che sembra sfuggire una definizione oggettiva. I nervi non mostrano dolore anche se, indubbiamente, sono coinvolti. Se indagiamo a fondo con strumenti sempre più sofisticati (per esempio con la fMRI, l'EEG o addirittura con misurazioni intracraniche) si trovano tante fenomeni e processi di grande interesse neurofisiologico e medico, ma non si trova alcunché richieda o implichi la coscienza. Ogni neurone riceve i suoi input e li manipola sulla base di reazioni chimiche totalmente sufficienti a descrivere il suo output. Non avviene niente di imprevisto. La combinazione dell'attività di miliardi di neuroni produce, a tempo debito, il comportamento e il linguaggio. In mezzo a loro però niente richiede qualcosa di aggiuntivo e quindi non c'è spazio per quel fenomeno misterioso che chiamiamo coscienza.

La coscienza è totalmente assente dalle osservazioni delle neuroscienze e al crescere del dettaglio nella nostra conoscenza del funzionamento del sistema nervoso, il problema della coscienza diventa progressivamente più difficile, anziché il contrario.

## La ricerca della coscienza: le principali teorie in gioco

Il campo della coscienza si trova in una situazione molto simile a quella delineata da Thomas Kuhn in quei casi in cui un paradigma viene messo in discussione (Kuhn 1962): in questo caso il paradigma dominante prevede che la coscienza sia un fenomeno speciale dentro il cervello. Ci sono molte analogie tra la situazione attuale delle scienze che cercano di spiegare la coscienza e altri casi famosi del passato, come l'elettricità, l'etere luminifero, il calorico, il flogisto, la precessione del perielio di Mercurio. In tutti questi casi vi erano tante ipotesi quanti i ricercatori coinvolti e i loro sforzi sembravano condannati a non cogliere mai l'essenza del problema.

Per esempio, nel loro recente articolo *Theories of Consciousness*, Anil Seth e Timothy Bayne elencano ben 22 ipotesi totalmente incompatibili sulla natura della coscienza (Seth e Bayne 2022). Questa proliferazione crea qualche imbarazzo nella comunità e si è così assistito a una progressiva decantazione del settore che ha portato alla selezione di quattro ipotesi: Neural Global Workspace Theories (NGWT), Integrated Information Theories (IIT), Higher-Order Theories (HOT) e Recurrent Process Theories (RPT) come descritto nell'ultimo articolo di Mariana Lenharo «[The consciousness wars: can scientists ever agree on how the mind works?](#)» (Lenharo 2024). L'elenco ha lasciato fuori, con qualche disappunto dei diretti interessati sia gli approcci basati sulla meccanica quantistica sia quelli che occhieggiano a varie forme di pansichismo; approcci che lascerà da parte in quanto richiedono delle sospette revisioni *ad hoc* della fisica in modo da poter aggiungere, di fatto, qualcosa di simile all'anima.

## La coscienza e l'informazione: una questione aperta

Le quattro possibili ipotesi hanno qualcosa in comune: sono basate sull'assunto secondo cui l'esperienza cosciente sarebbe una specie di "informazione a colori" che sarebbe generata all'interno di un sistema computazionale. Al di là dei nomi altisonanti dei quattro approcci, l'idea di base è sorprendentemente simile: la coscienza sarebbe un mondo virtuale generato, all'interno del



sistema nervoso, grazie all'elaborazione dell'informazione in qualche modo particolare. Se l'espressione «in qualche modo particolare» appare vaga, è perché lo è. Ognuna delle quattro ipotesi propone un modo diverso per passare dall'informazione invisibile e incosciente alla informazione *colorata* (ovvero l'esperienza fenomenica). Per esempio, secondo la IIT, originariamente proposta da Giulio Tononi nel 2004 (Tononi 2004), l'informazione dentro il cervello sarebbe elaborata in modo da essere integrata e, per questo motivo, diventerebbe coscienza. Per quale motivo l'informazione integrata dovrebbe essere fenomenicamente esperibile e da chi, Tononi non l'ha mai detto. Né ha mai detto qualche connessione dovrebbe esistere tra questo tipo di informazione e la selezione naturale. Se anche la sua ipotesi fosse vera, la connessione tra integrazione dell'informazione e selezione naturale sarebbe una fortunata coincidenza. Il che non pare particolarmente convincente.

Insomma, la teoria dell'IIT, nonostante sia stata al centro di numerose ricerche e abbia assorbito ingenti risorse, non ha mai spiegato la domanda di fondo: perché l'informazione integrata (che alla fine è espressa da equazioni che contengono solo termini quantitativi) debba diventare l'esperienza cosciente con le sue familiari proprietà? Per non parlare del fatto che la teoria soffre di problemi metodologici piuttosto seri tra i quali: i) non si sa (forse non è nemmeno possibile) misurare direttamente la presunta quantità di informazione integrata di un sistema concreto, ii) le quantità misurate dipendono da grandezze ipotetiche e non concretamente in atto, iii) le ipotesi fatte non hanno una base fisico-naturalistica ma sono (per stessa ammissione dei ricercatori) dei postulati, iv) non si capisce da un punto di vista funzionale-evoluzionistico per quale motivo questo tipo di struttura dovrebbe essere stato selezionato dal sistema nervoso. Ma anche se tutti questi problemi fossero risolti e si riuscisse a dare una base empirica solida all'informazione integrata rimarrebbe comunque la domanda di partenza: perché l'informazione integrata dovrebbe dar luogo alla coscienza? Non si sa.

Per questi motivi, un gruppo di 124 celebri scienziati e filosofi in un [articolo](#) che ha fatto scalpore ha accusato la IIT di non essere una vera ipotesi scientifica ma di essere un caso di pseudoscienza (Fleming, Frith et al. 2023). Ne è seguito un vero e proprio bailamme «La lettera provocò reazioni negative da parte di altri scienziati che ritenevano che un tale attacco potesse aggravare le divisioni e danneggiare la credibilità del campo. I firmatari riferirono di aver ricevuto e-mail minacciose contenenti minacce velate. Ricercatori da entrambi i lati dello spettro politico persero il sonno a causa di tweet accusatori. Alcuni persino contemplarono l'idea di abbandonare del tutto la scienza» (Lenharo 2023). Mi soffermo su una reazione significativa da parte dei ricercatori coinvolti: il fatto di considerare questo tipo di attacco come lesivo della credibilità del settore e quindi come inopportuno.

A me pare curioso che si debba valutare una contestazione fatta sul merito di una ipotesi e di un metodo sulla base della politica e della credibilità degli autori. Immaginiamo di tornare indietro al 1598 e di applicare lo stesso tipo di ragionamento: dovremmo forse chiedere di rispettare il sistema ibrido di Tycho Brae per non sminuire la autorità degli astronomi geocentrici danesi? Non si vede perché ... sarebbe una pericolosa ingerenza degli interessi delle persone all'interno di un settore scientifico e la direzione della ricerca stessa.

D'altronde, se l'IIT piange, le altre teorie non ridono. Gli altri tre gruppi citati (NGWT, HOT e RPT) non godono di ottima salute; sono teorie in circolazione da oltre una trentina d'anni e non sono mai state capaci di progredire oltre la formulazione originaria. Sostanzialmente si differenziano dalla IIT solo per il particolare tipo di computazione che sarebbe necessario per trasformare l'informazione in coscienza. Per la NGWT, il passaggio sarebbe provocato dal fatto che l'informazione *entrerebbe* all'interno di uno spazio di lavoro che metterebbe insieme più fonti per produrre un comportamento unificato. Ma perché questa *entrata* dentro lo spazio di lavoro dovrebbe

conferire una natura cosciente all'informazione non è assolutamente spiegato. Non viene neppure spiegato perché lo *spazio di lavoro* (il *global workspace*, appunto) dovrebbe essere cosciente. La HOT e la RPT propongono giustificazioni simili che non toccano il cuore del problema. Per la HOT il passaggio chiave sarebbe avere rappresentazioni e meta-rappresentazioni, mentre per la RPT le informazioni dovrebbe andare avanti e indietro dalle aree sensoriali a stazioni di elaborazioni successive. Per carità, sono tutte ipotesi legittime circa i meccanismi di elaborazione del sistema nervoso. Ma ci dicono qualcosa sulla domanda vera, almeno nel caso della coscienza, non pare proprio anche se questi approcci sono stati lautamente finanziati per decenni con l'obiettivo di scoprire le cosiddette basi neurali della coscienza.

## Il paradigma dominante negli studi sulla coscienza

L'impressione è che in tutti questi casi ci siano due meccanismi in gioco: uno che riguarda la sociologia del metodo scientifico e l'altro che riguarda quelle ipotesi nascoste che, dal buon senso comune, finiscono per condizionare inconsciamente la ricerca scientifica come ammonito da Albert Einstein (Einstein 1954) in un famoso passo introduttivo circa le ipotesi sulla natura dello spazio tempo: "Nell'interesse della scienza, è necessario impegnarsi ripetutamente nella critica di questi concetti fondamentali, affinché non siamo inconsapevolmente dominati da essi. Questo diventa evidente specialmente in quelle situazioni che coinvolgono lo sviluppo di idee in cui l'uso coerente dei concetti fondamentali tradizionali ci porta a paradossi difficili da risolvere."

Nel caso della coscienza si vede spesso in atto il gioco del credito epistemico: gruppi di ricercatori affermati in discipline prestigiose prendono a credito la propria autorevolezza per ottenere finanziamenti e risorse allo scopo di svolgere attività in campi nei quali non hanno alcuna competenza di fatto a meno di non presupporre, a priori, che il problema rientri nel loro campo di pertinenza. Da questo punto di vista, il principale motivo per cercare la coscienza dentro il sistema nervoso centrale non riguarda tanto i nostri neuroni, quanto il settore disciplinare dei neuroscienziati che ritengono che la coscienza debba essere risolta da loro.

Inoltre la ricerca sembra essere guidata (nel senso di condizionata definito da Albert Einstein nel passo precedente) l'idea di senso comune, ma non per questo particolarmente convincente, secondo cui la coscienza sarebbe all'interno del cervello. In tempi recenti, questa credenza popolare si è intrecciata con la teoria dell'informazione e con l'impressione che l'informazione sia l'analogo moderno dell'anima. L'informazione è concepita da molti come una specie di livello (di sostanza) distinta dai processi fisici in quanto tali che si sposterebbe dagli oggetti al sistema nervoso e da questo a vari dispositivi. Mi rendo conto moltissimi obietteranno che questa formulazione è troppo semplice e non rende giustizia alla teoria dell'informazione, ma se si consultano testi sia di neuroscienze che di intelligenza artificiale è difficile sottrarsi all'impressione che l'informazione venga trattata come qualcosa che «si genera», «si trasmette», si «interpreta», si «immagazzina», si «trasforma» e tante altre cose che ricordano da vicino l'idea di anima o di etere.

D'altronde tutte e quattro le classi di teorie citate (IIT, GNWT, HOT e RPT) hanno in comune l'ipotesi secondo cui dentro i sistemi nervosi ci sarebbe questa sostanza, chiamata informazione, che, nelle opportune condizioni, si trasformerebbe nella coscienza. Comprendo che molti obietteranno che non oggi nessuno pensa che l'informazione sia una sostanza e non potrei essere più d'accordo. Ma se non lo è, come può essere la base per diventare la nostra coscienza? E perché utilizzare una terminologia che implica che l'informazione sia un livello della realtà addizionale rispetto ai processi fisici e quindi implicitamente o inconsciamente avallando una specie di dualismo, senza avere il coraggio di usare questo termine?

Questa è la situazione in cui versa oggi la ricerca della coscienza. Tornando alla celebre analisi di Thomas Kuhn delle rivoluzioni (Kuhn 1962), possiamo tentare di delineare il paradigma degli studi della coscienza? È abbastanza facile e basta considerare le affermazioni comuni alla maggior parte dei ricercatori coinvolti. In sintesi, la struttura del paradigma dominante nel campo degli studi sulla coscienza è il seguente:

- la coscienza è fisica (questo di solito viene preceduto da una invettiva contro Cartesio ...);
- la coscienza è interna al sistema nervoso e ne è una proprietà;
- la coscienza non è osservabile direttamente perché altrimenti l'avremmo già vista;
- noi facciamo esperienza della coscienza che si trova all'interno del nostro sistema nervoso;
- la coscienza emerge o è identica dall'informazione contenuta all'interno del sistema nervoso che acquista proprietà speciali per qualche motivo misterioso (entrare nello spazio di lavoro, integrazione, connessioni rientranti, fenomeni quantistici). Chiamiamolo fattore X. Potremmo chiamare questo tentativo l'approccio del «primo passo» laddove il primo passo è proprio questo fattore X.

## L'approccio del "primo passo": un problema metodologico

Si vede subito come questo paradigma sia pieno di buchi e crei molti più misteri di quelli che risolve. In particolare, si vede subito come la sostituzione della coscienza con un misterioso fattore X non risolva, ma rimandi il problema e lasci aperta la domanda: perché il fattore X dovrebbe avere la capacità di creare la coscienza. Curiosamente, tutte le teorie citate sopra propongono un fattore X, ma nessuna spiega perché tale fattore dovrebbe essere tale nel creare la coscienza. Il «primo passo» non ci avvicina al problema della coscienza, più di quanto una scala non ci avvicini ad arrivare sulla Luna.

Da un punto di vista metodologico, l'approccio del «primo passo» è scorretto per un motivo molto semplice: sostituisce il vero problema con un problema apparentemente trattabile, ma diverso da quello della coscienza. In altri termini è come fare quello che fa l'ubriaco della barzelletta: cercare le chiavi sotto la luce del lampione e non dove le ha effettivamente perse. È una situazione analoga a quella nella quale descritta da Thomas Kuhn: la tendenza di un certo settore scientifico di considerare come scientifici solo quei problemi per i quali ha le risorse per affrontarli. Quando un problema, in questo caso la coscienza, esula dal perimetro dei problemi adatti a un certo settore, gli appartenenti a quel settore propongono altri problemi, interni al loro settore, e li promuovono a «primi passi» o altrettanti fattori X. Ovviamente non si avvicinano affatto alla soluzione del problema vero.

L'approccio del primo passo è profondamente scorretto perché sostituisce il vero problema con un altro e poi chiede di concentrarsi su questo secondo come se rappresentasse una soluzione. Faccio un esempio. Il problema vero è come sia possibile che qualcosa di fisico produca o sia responsabile per l'esperienza cosciente che, tutt'ora non trova una collocazione all'interno della descrizione fisica della realtà.

## LLM e coscienza: un legame possibile?

Sulla base di tutti questi elementi, possiamo trarre delle conclusioni utili per rispondere alla domanda di partenza: una LLM è vicina alla possibilità di creare una coscienza artificiale? La risposta è no in quanto, allo stato attuale degli studi della coscienza, non vi è alcuna conferma del fatto che la coscienza sia l'esito di processi computazionali interni. Gli LLM, che sono basati sulla tecnologia Transformer ovvero sulla sistematica estrazioni di probabilità condizionate, non hanno

alcun riferimento alla coscienza. Perché, come nel caso del fattore X suggerito dalle teorie (fallimentari fino a prova contraria) oggi in circolazione, la sistematica raccolta di probabilità condizionate dovrebbe trasformarsi in coscienza? Perché il trilione e più di probabilità condizionate che consentono a ChatGPT di manifestare una notevole competenza linguistica dovrebbe rivestirsi di esperienza fenomenica? È indubbio che gli esseri umani dotati di competenza linguistica sono coscienti, ma il legame tra il linguaggio e la coscienza non rientra in alcuna teoria o ipotesi circa la coscienza.

## Bibliografia

Chalmers, D. (2022). *Could large language models be conscious?* NeurIPS conference, New Orleans.

Chomsky, N., I. Roberts e J. Watumull (2023). «Noam Chomsky: The False Promise of ChatGPT» in *The New York Times*(March 23).

Einstein, A. (1954). Foreword.in (a cura di). M. Jammer *Concepts of Space*. Harvard, Harvard University Press.

Fleming, S., C. Frith, M. Goodale, H. Lau, J. E. LeDoux, A. L. Lee, M. Michel, A. Owen, M. A. Peters e H. A. Slagter (2023). «The Integrated Information Theory of Consciousness as Pseudoscience» in.

Kuhn, T. S. (1962). *La struttura delle rivoluzioni scientifiche*. Milano, Einaudi.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago, The University of Chicago Press.

Lenharo, M. (2023). «AI consciousness. Scientists say we urgently need answers» in *Nature*(21 December).

Lenharo, M. (2023). «Consciousness theory slammed as ‘pseudoscience’ — sparking uproar» in *Nature* September: 1-6.

Lenharo, M. (2024). «Consciousness the future of an embattled field» in *Nature* 625: 438-440.

Manzotti, R. e V. Tagliasco (2001). *Coscienza e Realtà. Una teoria della coscienza per costruttori e studiosi di menti e cervelli*. Bologna, Il Mulino.

Melloni, L., L. Mudrik, M. Pitts, K. Bendtz, O. Ferrante, U. Gorska, R. Hirschhorn, A. Khalaf, C. Kozma, A. Lepauvre, L. Liu, D. Mazumder, D. Richter, H. Zhou, H. Blumenfeld, M. Boly, D. Chalmers, S. Devore, F. Fallon, F. P. de Lange, O. Jensen, G. Kreiman, H. Luo, T. I. Panagiotaropoulos, S. Dehaene, C. Koch e G. Tononi (2023). «An adversarial collaboration protocol for testing contrasting predictions of global neuronal workspace and integrated information theory» in *PLoS One* 18(2): e0268577.

Pezzulo, G., T. Parr e K. J. Friston (2024). «Active inference as a theory of sentient behavior» in *Biol Psychol*: 108741.

Seth, A. e T. Bayne (2022). «Theories of consciousness» in *Nature Reviews Neuroscience* 23(7): 439-452.

Tononi, G. (2004). «An information integration theory of consciousness» in *BMC Neuroscience* 5: 1-22.

Turing, A. M. (1950). «Computing Machinery and Intelligence» in *Mind* 59: 433-460.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser e L. Polosukhin (2017). «Attention Is All You Need» in *arXiv*: 1706.03762v5



# Autopsia psicologica dell'autore di reato: un modello informatico per decifrare il movente

L'autopsia psicologica emerge come un nuovo strumento di indagine sulle motivazioni che guidano i criminali. Unendo metodologie tradizionali con innovative tecniche d'indagine, l'autopsia psicologica può offrire una nuova prospettiva nella comprensione dei comportamenti devianti

Di **Vincenzo Lusa**, Fellow American Academy of Forensic Sciences-Università Sapienza (Roma)

## Abstract

Scopo del presente lavoro consiste nell'accertare, nel soggetto attivo di reato, le reali motivazioni di natura psicologica (ovvero di natura psichica) da porre alla base delle commissioni di atti criminali. Per conseguire il suddetto risultato, l'autore propone la creazione di un modello informatico che riecheggi le modalità parametriche proprie dell'"autopsia psicologica". Quest'ultima, com'è noto, rappresenta una tipologia di accertamento che è effettuato *post-mortem* sulla vittima di reato, ponendo in luce il vissuto della persona deceduta, la sua personalità e le relazioni interpersonali che la persona offesa dal reato aveva attuato con vari soggetti e forse anche con il suo omicida. Quanto espresso, da attuarsi mediante una raccolta d'informazioni recepite da persone in contatto con la vittima, le quali possano fornire utili elementi al fine di ricostruire il suddetto quadro. Ebbene, agendo in maniera inversa, cioè utilizzando l'autopsia psicologica sul reo, invece che sulla sua vittima, attraverso l'esame dell'*iter criminis*, in altre parole del percorso psichico dal quale si origina il reato, si potrebbero conseguire due risultati: il primo connesso alla valutazione del proposito criminoso mediante la comprensione del "movente", con quest'ultimo termine s'intende la causa psichica che spinge il soggetto a compiere l'azione criminosa, e del "dolo di azione" (intenzione, volontà e rappresentazione dell'evento) che la dottrina attuale ritiene, erroneamente a parere di quest'autore, distinti tra loro. Il secondo risultato è invece pertinente all'analisi del suddetto movente dal punto di vista della capacità di intendere e di volere: ovverosia se il movente sia viziato da patologie o disturbi di natura psichica o psichiatrica. Ricostruendo i predetti parametri, unificandoli e inserendoli in un modello informatico gestito da un algoritmo, ebbene si potrebbe giungere a comprendere l'effettiva personalità del reo e le motivazioni poste alle fondamenta del suo agire criminale.

## Introduzione

Come è noto, il quotidiano impiego dell'intelligenza artificiale, in tutti i campi dello scibile e delle tecnologie umane, sta divenendo sempre più consueto da parte degli operatori addetti ai settori industriali e del sapere; intesa quest'ultima accezione come l'indagine di natura scientifica sulla natura. A questa considerazione non si sottrarre il settore forense e più in particolare quello deputato alla ricerca degli autori di reato, anche con specifico riferimento alla prevenzione del crimine e alla comprensione delle motivazioni di natura psicologica poste alla base dell'agire criminale. Negli

ultimi anni invero si è originato il settore scientifico/criminologico volto alla comprensione degli eventi criminosi comprendendo in quest'ultimo alveo gli omicidi seriali e gli atti delittuosi di natura particolarmente efferata. Il summenzionato settore criminologico è stato definito con il termine di "Analisi criminale". Con la predetta terminologia, Aramini nel 2002 (1) ha inteso quel processo metodologico finalizzato appunto allo studio di eventi criminosi, indagando sul probabile omicida, sul reato da quest'ultimo commesso e il luogo ove l'atto criminoso si è verificato. Tale metodologia non risponde solamente ai fini dell'individuazione del reo ma essa dispiega le sue potenzialità anche nell'ottica della prevenzione di reati. Oggigiorno, diversi algoritmi sono utilizzati nel campo del *digital forensic*, ed essi hanno per oggetto lo studio delle dinamiche (*modus operandi*) mediate le quali un reato si è compiuto al fine di redigere il profilo psicologico del presunto criminale. Questa tipologia di analisi criminale si basa su uno stampo psicologico dinamico chiamato *profiling*. Pertanto, analizzando a livello informatico il *modus operandi* dell'ipotetico colpevole in relazione alla sua vittima e alla scena del crimine si potrebbe risalire al movente del presunto reo. Turvey (2) ad esempio ha posto l'ipotesi che l'autore del reato utilizzi un determinato modello di comportamento nel momento in cui commette un crimine e che tali modalità attuative della condotta si evincano anche sulla scena del crimine. Circa poi quest'ultimo scenario, ovvero la scena del crimine, unendo il profilo criminologico dell'autore di reato ai luoghi ove i delitti sono attuati si riesce a prevenire reati come quelli contro il patrimonio, utilizzando a tale scopo il software GIS (Geographic Information System) che genera mappe geo-referenziate e pertinenti alla ripartizione dei reati al fine di individuarne l'autore. Ai fini della presente ricerca, ci focalizzeremo esclusivamente sul profilo psicologico del criminale al fine di ricavarne il movente (ovvero l'impulso psicologico alla commissione del reato) mediante la ricerca delle modalità parametriche proprie dell'autopsia psicologica attuata sul reo attraverso l'esame dell'*iter* criminis, ovvero del percorso psichico dal quale si origina il reato comprendendo così le motivazioni poste alla base dell'agire criminale. Procedendo in tal modo, saranno così rilevati vari parametri di natura psicologica individuati nel soggetto attivo di reato, parametri che in futuro potrebbero essere inseriti in un modello digitale informatico con lo scopo di evincere non solo la personalità del reo (o presunto che sia) ma soprattutto il "perché" costui si sia determinato ad agire compiendo reati contro la persona. Invero, qualsiasi modello informatico di natura forense che sia basato su algoritmi gestiti da un'intelligenza artificiale abbisogna di parametri connotati da univoca certezza. Lumeggiante esempio circa quanto narrato, lo rinveniamo quando la giurisprudenza erroneamente continua a differenziare il dolo dal movente del reato generando così un'alea d'incertezza tale da rendere un modello informatico non in grado di percepire le reali motivazioni dalle quali si sono originate le azioni delittuose attuate dal reo.

## Materiali e metodi

Come delucidato nell'introduzione al presente lavoro, al fine di creare un modello digitale informatico atto a definire sia la personalità del reo, sia il movente che lo ha determinato ad agire in contrasto con le regole codificate nel codice penale, appare necessario rilevare i parametri di natura giuridica nonché psicologica già in parte presenti nel codice penale. Una volta conseguiti i predetti parametri occorrerà successivamente porli in sistema con le Scienze forensi. Il predetto ambito scientifico è, come noto, deputato alla comprensione della personalità dell'*offender* e del suo *network* cerebrale ove l'ideazione e la risoluzione del piano criminale si sono originati e poi evoluti manifestandosi nell'ambiente mediante l'azione.

- (1) Il Dolo di azione. Di primo acchito, occorrerà quindi prendere in esame l'elemento psicologico del reato espresso tramite l'istituto giuridico penalistico del "Dolo di azione". Conseguentemente, nel nostro codice penale l'elemento soggettivo del reato è scandito agli articoli 42 e 43 che, com'è noto, si traducono negli istituti giuridici del Dolo e della Colpa. All'interno dell'articolo 42 del c.p. compare peraltro anche l'espressione "con coscienza e

volontà” riferita alla commissione di un’azione o omissione identificabile come reato. Tal espressione, come potremo appurare appare criticabile ai nostri fini. Pertanto: l’articolo 42 codice penale recita che: “Nessuno può essere punito per un’azione od omissione preveduta dalla legge come reato, se non l’ha commessa con coscienza e volontà. Nessuno può essere punito per un fatto preveduto dalla legge come delitto, se non l’ha commesso con dolo, salvo i casi di delitto preterintenzionale o colposo espressamente preveduti dalla legge. La legge determina i casi nei quali l’evento è posto altrimenti a carico dell’agente, come conseguenza della sua azione od omissione. Nelle contravvenzioni ciascuno risponde della propria azione od omissione cosciente e volontaria, sia essa dolosa o colposa.” Di seguito l’articolo 43 codice penale definisce che: “Il delitto: è doloso, o secondo l’intenzione, quando l’evento dannoso o pericoloso, che è il risultato dell’azione od omissione e da cui la legge fa dipendere l’esistenza del delitto, è dall’agente preveduto e voluto come conseguenza della propria azione od omissione; è preterintenzionale, o oltre l’intenzione, quando dall’azione od omissione deriva un evento dannoso o pericoloso più grave di quello voluto dall’agente; è colposo, o contro l’intenzione, quando l’evento, anche se preveduto, non è voluto dall’agente e si verifica a causa di negligenza imprudenza o imperizia, ovvero per inosservanza di leggi, regolamenti, ordini o discipline “. Secondo una definizione usualmente accettato in dottrina a cura dell’Antolisei (3), il dolo rappresenta tout court la cattiva volontà dell’autore del reato. Il dolo si compone di una rappresentazione che consiste nella visione anticipata del fatto-reato e di una risoluzione, che in realtà si traduce in una manifestazione del volere finalizzata a compiere il fatto: ovverosia l’evento deve essere voluto dal soggetto. In tal guisa, si può discorrere di un momento conoscitivo del dolo, da parte dell’agente, che di norma precede quello volitivo. Il suddetto momento conoscitivo, secondo l’Antolisei (3), è determinato dalla piena consapevolezza dell’illiceità del gesto che si compie e dei contenuti violati propri della norma penale che s’infrange. In altri termini l’agente, con la sua cattiva volontà, non solo ha attuato un fatto previsto dalla legge come reato, ma lo ha anche fortemente voluto. E invero, quanto sopra espresso, chiarisce il perché nel codice penale si parla di coscienza dell’illiceità dell’azione, poiché è proprio nell’articolo 42 del codice penale, che si rinviene, al secondo comma, il precetto volto a indicare che è proprio nel dolo che risiede il normale criterio di volontà soggettiva per i delitti. La definizione dei summenzionati precetti è insita nel successivo articolo 43 del codice penale ove si può evincere l’architettura giuridica sulla quale il dolo si sorregge. Tale struttura è caratterizzata da un’intenzione, una previsione e la volontà di realizzare il fatto. La giurisprudenza tuttavia interpreta l’intenzione e la previsione inglobandola nella rappresentazione alla quale fa seguito la volontà dell’evento che si causa. In effetti, con l’accezione: “rappresentazione” s’intende il momento intellettuale del dolo ovvero la visione anticipata di tutti i momenti salienti del fatto-reato. Con la fase volitiva invece s’intende che la volontà dell’agente sia finalizzata all’effettiva realizzazione della condotta. Si discorre ancora in dottrina (3), circa la condotta che l’agente serba al momento della commissione del fatto; se essa invero si fonda su un’azione che deve essere cosciente e volontaria e diretta alla realizzazione dell’evento e connessa a quest’ultimo per il tramite di un nesso di causalità. La risposta certo appare affermativa e su quest’aspetto anche noi ne conveniamo. Comunque, e con ciò torniamo alle osservazioni in precedenza compiute, è sul parametro giuridico inerente alla previsione dell’evento che sorgono le maggiori perplessità perché, come vedremo, la giurisprudenza, non a torto, distingue il dolo dalla coscienza, quest’ultima intesa come apparato psichico e mentale del soggetto da inquadrarsi nel più ampio spettro possibile della scienza dedicata allo studio sull’encefalo oggi consentita dalla rappresentazione che il soggetto effettivamente possiede dell’evento. Infatti, occorre approfondire se con il termine “previsione” si voglia intendere la conoscenza da parte del soggetto delle conseguenze dell’evento (tuttavia alla luce delle Neuroscienze forensi ciò pare esulare dalle potenzialità che il dolo in realtà esprime). Alla luce delle considerazioni



sopra esposte si possono rilevare i primi parametri psicologici posti alla base della commissione dei reati. Di conseguenza i reati sono compiuti in virtù dei seguenti parametri:

- a) coscienza e volontà dell'azione (così detta *suitas*);
- b) secondo l'intenzione;
- c) l'evento dannoso è dall'agente preveduto e voluto.

d) Imputabilità. Ai sensi dell'art. 85 c.p.: «nessuno può essere punito per un fatto preveduto dalla legge come reato se, al momento in cui lo ha commesso non era imputabile. E' imputabile chi ha la capacità di intendere e di volere». Pertanto è imputabile chi annovera a sé sia la capacità di “intendere” che quella di “volere” si riverberano sulla colpevolezza. Viceversa, per non essere imputabili occorre che vengano meno o la capacità di intendere o quella di volere; ovvero ambedue. Si può essere dichiarati non imputabili per cause connesse allo stato tossicologico del soggetto (ubriachezza o intossicazione provocata da vari agenti nocivi), ovvero per motivi fisiologici (minore età) o psicologici dell'individuo (infermità di mente o sordomutismo). La capacità di intendere è concepibile come la percezione volta a comprendere la realtà esterna e di conseguenza cogliere il disvalore delle proprie azioni o omissioni. In altri termini sono consapevoli che sto violando le norme contenute nel codice penale poiché la legge penale, *ex art. 5 del c.p.*, non ammette ignoranza. La capacità di volere è riferita al controllo degli impulsi cerebrali; questi ultimi, sotto l'egida della volontà, conducono o no l'individuo a agire. Dove entrambe le condizioni siano presenti, allora potrà trovare applicazione la sanzione penale; quest'ultima invero sarebbe inapplicabile a colui il quale che non sia in grado di “intendere” e di “volere”. Di conseguenza i reati altresì sono attuati in base alla: capacità di “intendere” e capacità di “volere”

Alla luce delle considerazioni sopra esposte si evincono gli ulteriori due parametri psicologici posti alla base della commissione dei reati:

- e) capacità di “intendere”;
- f) capacità di “volere”.

3) Coscienza e volontà dell'azione (*suitas*) e l'imputabilità. La coscienza e la volontà dell'azione sono da considerarsi distinte dalla capacità di intendere e di volere. Si può proporre, a lusinghioso esempio di quanto sopra espresso, il caso di colui che in preda a follia possiede sì, la coscienza (in quanto sveglio e cosciente) e la volontà (come movimento della massa muscolare determinata dagli impulsi mentali) di uccidere, ma di certo non annovera a sé la capacità di comprendere il disvalore sociale delle azioni che compie e che in un momento di sanità mentale certo non avrebbe desiderio di realizzare: di conseguenza difettano la capacità di intendere e di volere (4). Pertanto, in virtù delle considerazioni stilate, occorrerà valutare se sotto il dominio del controllo volontario possono ricadere i così detti “atti automatici” che dottrina di settore come l'Antolisei (3) suddivide in “istintivi”, “riflessi” ed “abituati”. Con la prima categoria s'intendono quella tipologia di atti che il corpo umano, sotto l'impulso mentale, attua in determinate circostanze ad es. coprirsi il volto quando qualcosa sopraggiunge repentinamente verso di esso. Con gli “atti riflessi” si circoscrivono invece quegli atti che si pongono in essere senza che la mente possa esercitare su di essi alcuna signoria di volere come ad esempio la respirazione o il battito cardiaco; ed infine si parla di “atti abituali” riferendosi a quelle movenze che un soggetto compie proprio sotto l'influsso dell'abitudine in modalità automatica, come ad es. il fumatore incallito che getta il mozzicone di una sigaretta in un bosco provocando un incendio. Ovvero un atto delittuoso compiuto in stato di dormiveglia. La *suitas* quindi individua il limes oltre il quale la condotta del soggetto non diviene più attribuibile all'agente, giacché costui non annovera più il dominio di se stesso. Appare palese che, circa i suddetti “atti automatici” ebbene, esclusa la prima e la seconda categoria tutte le altre tipologie di atti ricadono sotto la signoria del

volere e sono pertanto evitabili attuando uno sforzo di volontà ed in effetti, anche per la prima categoria (ovvero gli “istintivi”) la spinta mentale potrebbe, ma non sempre, vincere l’istinto e comportarsi diversamente, anche se con ovvie difficoltà. La “coscienza e volontà” dell’azione ancora una volta si pone su un piano distinto dall’imputabilità. Precisando dunque quanto sopra riportato, la Cassazione penale ha stigmatizzato recentemente il principio volto ad affermare che l’imputabilità, quale capacità di intendere e di volere, e la colpevolezza, quale coscienza e volontà dell’atto illecito, agiscono su piani che sono ben distinti tra loro, benché la prima, come substrato naturalistico della responsabilità penale, dovrà essere accertata in modalità prioritaria rispetto alla seconda (così la Cass. Pen, Sez. VI, n.41083 del 4 ottobre 2013). Conseguentemente il parametro: a) “coscienza e volontà dell’azione” è distinto dai parametri: d) capacità di “intendere” e) capacità di “volere”.

- 4) Il movente. Secondo la giurisprudenza imperante, il movente è la causa psichica della condotta umana; esso costituisce lo stimolo che ha indotto l’individuo ad agire (cfr. sentenza Cass. n. 466/1993 Sez. 1, dep. 1994, Ha., Rv. 196106). Sempre secondo la suddetta cassazione penale il movente si distingue dal dolo che, come già delineato, è da considerarsi come elemento costitutivo del reato e che riguarda la sfera della rappresentazione e volizione dell’evento. Di conseguenza il “mero movente” precisa la Corte pur potendo contribuire all’accertamento del dolo, costituendo una potenziale circostanza inferenziale, non coincide con il dolo del quale può rappresentare, invece, il “presupposto”. A nostro avviso, e al fine della costruzione del modello informatico basato su parametri certi e univoci, ciò appare censurabile per le motivazioni che si espongono anche alla luce delle Neuroscienze forensi. Il termine “Neuroscienza” compare all’incirca negli anni 1970, esso si prefigge di comprendere gli studi di settore inerenti all’indagine e valutazione del sistema nervoso esaminato dal punto di vista anatomico, genetico e funzionale. In esito al suddetto termine Negli ultimi anni, negli ambienti forensi e non solo in questi ultimi, a lungo si è discusso intorno a quello che è stato chiamato “*The warrior gene*” ovvero il gene guerriero (MAO) attribuendo a quest’ultimo ruolo e funzioni non del tutto pertinenti all’unità biologica in questione e ciò dal punto di vista della corretta scientificità dell’argomento in questione (4). Le monoamine rappresentano un gruppo di neurotrasmettitori, i quali annoverano a sé serotonina, dopamina e noradrenalina; questi ultimi sono sintetizzati tramite azione di specifici enzimi come le monoamine ossidasi (MAO) e catecolo O- metiltransferasi (COMT). In seguito al rilascio delle monoamine a livello delle sinapsi, queste rappresentano le autostrade delle vie di comunicazione tra i neuroni, determinati trasportatori sono addetti alla loro ricettazione a livello cellulare allo scopo di interrompere il segnale neuronale. Le monoamine ossidasi sono presenti sul cromosoma femminile X. Le monoamine ossidasi possono divenire oggetto di anomalia genetica, nel senso che i geni, i quali causano la codificazione delle monoamine ossidasi, che come si è detto sono presenti sul cromosoma X femminile, possono andare incontro ad alterazione genetica e ciò si traduce nella presenza di un allele più corto, ovvero una bassa espressione dell’enzima MAO-A che diviene quindi MAOA-L (L=low). Le monoamine ossidasi sono ampiamente diffuse a livello nervoso e sono associate all’attuazione di comportamenti impulsivo e aggressivi (5). La suddetta mutazione genetica dell’enzima MAO-A, come versione polimorfica MAOA-L nella variante corta, ne comporta una produzione limitata tale da provocare, nei soggetti portatori di tale mutazione, un’inclinazione all’aggressività o atti impulsivi tale da sfociare in manifestazioni antisociali di natura violenta. Da qui la denominazione del MAOA-L come gene del “guerriero”, nomignolo peraltro attribuito per la prima volta al suddetto polimorfismo dalla giornalista Ann Gibbons, in occasione di un *meeting* scientifico dell’AAPA (American Association of Physical Anthropology) quando furono discusse,

in sede di conferenza, le cause dell'elevato tasso di aggressività posseduto dai guerrieri Maori (conosciuti come i guerrieri senza paura) rispetto ad altre popolazioni locali, attribuendo la suddetta violenza alla presenza, nell'assetto genetico dei Maori, dell'allele MAOA-L. Caspi peraltro, nell'anno 2002 (6) effettuò un importante studio su un vasto campione di cittadini neozelandesi di sesso maschile monitorati dall'infanzia sino ai 26 anni di vita. Lo studio evidenziò che i soggetti in possesso di una variante del MAOA a bassa attività (Low) e cresciuti in infanzia in un ambiente maltrattante svilupparono - rispetto a individui egualmente cresciuti in ambiente sfavorevole, ma dotati al contrario della variante in MAOA ad alta attività (HIG) - un rischio superiore di manifestare comportamenti antisociali. Questo denota che i soggetti che siano inconsapevoli portatori della variante Low, e che si trovino a operare in ambiente caratterizzato da attività legata a una forte componente di origine stressante, sono più inclini di altri a commettere manifestazioni violente. Il sistema serotonergico è ritenuto uno dei principali regolatori dei comportamenti impulsivo-aggressivi a livello cerebrale e in particolare esso è responsabile dell'aggressività proattiva (offensiva). Il sistema serotonergico è influenzabile dall'ambiente poiché eventi ad alta rilevanza di stress che sono occorsi in età precoce sembrano in grado di mutare la risposta serotonergica nell'adulto è quindi, per la finalità del presente lavoro, appare d'indubbia importanza il tenerne conto. In particolare, la serotonina (5-HT) è un neurotrasmettitore che principalmente agisce come freno inibitorio naturale poiché regola l'impulsività. Il basso livello di serotonina può determinare l'insorgere di atti antisociali (7). Inoltre il gene SCL6A4, che codifica per il trasportatore della serotonina (SERT), è un modulatore fondamentale della trasmissione serotonergica ed è in grado di ridurre i comportamenti aggressivi nell'uomo (8). Il riscontro di una riduzione dell'attività serotonergica a livello cerebrale può comportare un aumento del comportamento impulsivo e aggressivo ed esso stimola la corteccia frontale, zona del cervello che regola l'aggressività. Il polimorfismo "s" del gene SCL6A4 comporta una maggiore incapacità di adattamento alla presenza di condizioni ambientali sfavorevoli. È stato invero riscontrato che i portatori della variante s sono a forte rischio di depressione e attuano comportamenti suicidari o aggressivi. Il sistema dopaminergico è coinvolto nella genesi del *reward* che annovera a sé i comportamenti rivolti alla sopravvivenza di una specie, tra i quali ricerca di cibo, attività sessuale e comportamenti aggressivi (9). Il tutto finalizzato a stabilire il dominio nei confronti di un nemico; l'effetto *reward*, peraltro, procura nel soggetto che lo subisce una sensazione di appagamento. Il riscontro di una ridotta sensibilità nel sistema dopaminergico, che è finalizzato al controllo dei *reward*, può essere implicato nell'accrescimento patologico dell'aggressività. I geni maggiormente coinvolti nell'effetto modulante sull'attività del sistema dopaminergico sono SLC6A3 (trasportatore della dopamina) e COMT che è un enzima che prende parte alla regolazione dei livelli di dopamina a livello sinaptico ed è modulatore dell'aggressività. In particolare, la presenza in un soggetto dell'allele 10 del gene SLC6A3 (allele 10 poiché esso è composto da 10 ripetizioni di una sequenza di nucleotidi) è stato correlato a comportamenti impulsivo-aggressivi di natura estremamente violenta e comportamenti antisociali. Alla luce delle osservazioni di natura neuroscientifica forense, si deduce che il movente può essere inficiato, oltre che da patologie mentali, anche da disfunzioni biologiche tali da rendere l'individuo non libero ma determinato nel suo agire da anomalie di natura genetica. Deriva da quanto sin ad ora esposto che il "movente" del reato si origina in virtù dell'elaborazione del pensiero e che tale elaborazione può essere di natura sana o insana (si legga patologie mentali) e conseguentemente influenzato dalla genetica comportamentale individuale. Il movente in realtà, come si andrà a dimostrare, non è scisso dal dolo ma anzi esso è da porsi alla base dello stesso. In effetti già una parte della giurisprudenza larvamente già lo ha ammesso e in effetti: "Ne discende che

il dolo specifico non è che un movente normativamente qualificato, che si colloca al di là della coscienza e volontà del fatto”. (Corte di Cassazione - Sezioni Unite - 12 ottobre 2023 N. 41570). Parametro f) movente.

- 5) Iter criminis. l'iter criminis rappresenta la nuova elaborazione dell'elemento soggettivo del reato e del principio di colpevolezza e la subordinazione del dolo a tale nuova impostazione. Il dolo, di per sé, non appare più sufficiente, nella sua architettura giuridica, a rappresentare l'integralità della volontà dell'autore di reato ma, tuttalpiù, esso ne palesa l'effetto della volontà stessa, lasciando all'iter criminis il compito di riprodurre la complessa intelaiatura dell'ideazione e attuazione dell'intento criminale. In via preliminare, cerchiamo di chiarire cosa davvero s'intenda con iter criminis? L'antico brocardo latino cogitationis poenam nemo patitur (ovvero: nessuno può subire una pena semplicemente per avere pensato qualcosa) costituisce il cuore del cosiddetto principio di materialità del Diritto Penale. In effetti, non si potrà mai configurare un reato che conduca all'erogazione di una pena, se la volontà criminosa non si materializza in un comportamento esterno (invero alcuno potrà mai essere accusato formalmente di un omicidio semplicemente per averlo pensato). Invero, nell'art.1 del codice penale è prevista che per l'applicabilità di una sanzione penale occorre fare esplicito riferimento a "un fatto", effettivamente occorso, e che sia stigmatizzato dalla legge come reato. Conseguentemente, possiamo pervenire al seguente asserto: la mera meditazione, il pensiero insomma, non è assolutamente sufficiente al fine di ottenere un giudizio di condanna, ma occorrerà qualcosa di più affinché si possa ottenere una pena per avere causato un fatto antiggiuridico. Ora esaminiamo di quali fasi si compone l'iter criminis. Di seguito si elencano le singole frazioni: 1) "Ideazione", che nel reo agente rappresenta il canovaccio di matrice psicologica (del pensiero volto a un determinato fine) sul quale si fonda il reato di natura ovviamente dolosa. Con l'ideazione, infatti, l'agente si rappresenta nella sua psiche l'idea criminale che vuole perpetrare e che poi si cristallizza con la risoluzione della stessa; 2) "Preparazione" ove l'agente predispone ogni cosa, ove il tipo di reato preveda tale attività (si pensi agli atti criminali con finalità terroristiche); 3) "Esecuzione" ove l'agente mette in atto la condotta richiesta affinché il reato si manifesti; 4) "Perfezione" ove si realizza l'evento criminoso programmato, ovvero sia esso si presenta completo di tutti gli elementi che la stessa fattispecie delittuosa prevede affinché il reato possa ritenersi conforme ex lege; 5) "Consumazione" "quando il reato programmato e già "perfezionato" perviene alla sua massima gravità o in modalità coincidente o disgiunta (10). Il dolo di per sé non sarà sufficiente al fine di rappresentare l'essenza della volontà del reo, o presunto che sia. Viepiù, il movente che è ciò che muove l'agente ad agire può essere condizionato da altri parametri; tra questi ultimi alcuni di natura nosografica (pertinenti per esempio al campo delle patologie mentali) altri di natura biologica e quindi genetica. Come infine potremmo appurare, il frazionamento delle condotte e della volontà criminale, attuata in base allo schema suggerito dalla struttura giuridica dell'iter criminis, ci consente di analizzare la mens rea del soggetto di reato al fine di pervenire alla sua affermazione di colpevolezza, nella gradazione giuridica che la legge ci consente, e ciò mediante l'applicazione dei canoni del codice penale; ovvero alla sua esclusione sia per motivi nosografici, o in ordine alla accertamento di un'intenzione, da parte del prevenuto, che si chiarisca come manifestamente contraria al reato ascritto. Come agevolmente si evince, lo studio dell'iter criminis rappresenta il fulcro del problema che ci occupa in questa sede, poiché esso, come si è detto, nasconde in sé la chiave di volta per comprendere la natura del proposito criminale nonché la sua potenzialità. Il proposito si realizzerà come vedremo mediante l'ausilio della volontà. Il tutto si traduce nell'affermazione di colpevolezza da porsi carico di un soggetto come attribuzione del fatto-reato alla sfera psicologica di colui che l'ha commesso. Ma non solo. Al fine di pervenire all'esatta diagnosi dell'iter

criminis, Per poi quanto concerne quest'ultimo essenziale aspetto, vale la pena di rilevare che l'iter criminis, come conoscenza profonda del pensiero criminale, non deve essere limitato alla semplice "ideazione", così come riportata nella sopra declinata suddivisione, riguardo al tentativo di reato, ma essa si espande sino all'ultima frazione dell'iter in parola e quindi nella fase dedicata alla "consumazione" del reato mediante la sopravvivenza del "proposito" da individuarsi sino all'ultimo istante del completamento dell'azione criminale. Il "proposito", invero, sorregge l'azione criminale sino al suo compimento e ne cristallizza l'"intento". Ne discende che l'esatta configurazione del reato, come esaustività di tutti gli elementi che la stessa fattispecie delittuosa prevede, può sorreggersi sia su un proposito intenzionalmente razionale e quindi basato sulla piena capacità di intendere e di volere, ex articolo 85 c.p., che su un proposito di origine patologica e quindi rappresentare un vizio totale ovvero parziale di mente. In particolare, l'iter criminis pone in luce la grande differenza che sussiste tra l'"intento" e "il proposito". Quest'ultimo invero rappresenta la naturale escalation dell'"intento". Il "proposito" è da considerarsi come la consacrazione del mero "intento", poiché sul primo giace l'intera architettura del processo mentale del reo (sana o insana che essa sia) così da originare il compiuto delitto ovvero il suo tentativo. Ancora si può quindi affermare in conclusione che una volta definita la reale natura del "proposito" esso sarà poi attuato mediante la volontà del soggetto.

Parametri: g) Ideazione h) preparazione i) esecuzione l) consumazione m) perfezione.

## Risultati

Come esplicitato in premessa scopo del presente lavoro è finalizzato a comprendere quali siano le reali motivazioni di natura psicologica da porre alla base delle commissioni di atti criminali nel soggetto attivo di reato. Come la genetica comportamentale ci ha evidenziato l'effettiva cognizione del *network* cerebrale del soggetto attivo del reato è da considerarsi in ottica multifattoriale al fine di comprendere: (a) il movente del soggetto, sano o insano che sia, (b) la maturazione del "proposito" criminale, sano o in sano che sia. Tutto ci conduce alla creazione di un adeguato *iter criminis*, il quale *iter* si origina nella sfera psichica del reo o presunto reo. In virtù dei parametri sino ad ora evidenziati otterremo che il reato si basa e si manifesta mediante: a) coscienza e volontà dell'azione (così detta *suitas*); b) secondo l'intenzione, c) l'evento dannoso è dall'agente preveduto e voluto, d) capacità di "intendere"; e) capacità di "volere", f) movente, g) ideazione, h) preparazione, i) esecuzione, l) consumazione, m) perfezione.

Successivamente ricaviamo il seguente schema assimilabile a una *flow chart* da porre alla base del programma informatico gestito da un algoritmo. Lo schema che si propone include in se tutti i parametri sino ad ora esposti unificandoli. Di conseguenza, la motivazione - ovvero il movente di reato - si dedurrà in base allo schema in prosieguo esposto e che si dispiega in base ai parametri da noi identificati dal parametro g) al parametro m). Di conseguenza otterremo il seguente schema (10):

1) "Ideazione", che nel reo agente rappresenta il canovaccio di matrice psicologica (del pensiero volto a un determinato fine) sul quale si fonda il reato di natura ovviamente dolosa. Con l'ideazione (si legga intento) l'agente si rappresenta nella sua psiche l'idea criminale che vuole perpetrare e che poi si cristallizza con la risoluzione della stessa. In essa sono già contenuti i parametri riferiti al "dolo" ovvero: a) coscienza e volontà

dell'azione (così detta *suitas*); b) secondo l'intenzione, c) l'evento dannoso è dall'agente preveduto e voluto, d) capacità di "intendere"; e) capacità di "volere" (queste ultime due da accertare a livello psicologico nonché psichiatrico); f) movente;

2) "Preparazione" ove l'agente predispone ogni cosa, ove il tipo di reato preveda tale attività;

3) "Esecuzione" ove l'agente mette in atto la condotta richiesta affinché il reato si manifesti;

4) "Perfezione" ove si realizza l'evento criminoso programmato, ovvero sia esso si presenta completo di tutti gli elementi che la stessa fattispecie delittuosa prevede affinché il reato possa ritenersi conforme *ex lege*;

5) "Consumazione" "quando il reato programmato e già "perfezionato" perviene alla sua massima gravità.

L'evoluzione dal mero "intento" (criminale) alla sua definitiva cristallizzazione evolutasi nel "proposito" (criminale) si rinviene nei passaggi scanditi dal punto n.2 al punto n. 5. Infatti il "proposito" sorreggerà l'intera azione originatasi a livello mentale psicologico dall'"intento" sino alla manifestazione del reato che è definito appunto come "evento di reato"(11). I dati che dovranno essere inseriti nel suesposto schema si evinceranno in virtù dei risultati scaturiti dall'espletamento delle indagini giudiziarie e dalle perizie e consulenze di natura psicologica e psichiatrica nonché genetica d'attuarsi sul presunto colpevole. Agendo in base al declinato schema si otterrà la genuina natura del movente posto alla base del reato.

## Conclusioni

Il percorso argomentativo sino ad ora intrapreso ha dimostrato che è teoricamente possibile generare un modello informatico basato sul percorso motivazionale - si legga movente- serbato dal reo nell'atto di porre in essere un atto *contra legem* (12). Per conseguire il suddetto risultato, l'autore ha pertanto proposto di porre alla base del summenzionato modello l'iter criminis sul quale si fonda la genesi del pensiero criminale e imperniato sulle modalità parametriche proprie dell'autopsia psicologica tale da definire l'evoluzione dell'"intento" nefasto sino alla sua definizione del "proposito" "di natura criminale. L'iter criminis dovrà allora fondarsi sui parametri che abbiamo evinto, ovvero: a) coscienza e volontà dell'azione (così detta *suitas*); b) secondo l'intenzione, c) evento dannoso preveduto e voluto dall'agente, d) capacità di "intendere", e) capacità di "volere", f) movente, g) ideazione, h) preparazione, i) esecuzione, l) consumazione, m) perfezione. In particolare, i parametri scanditi dalla lettera g) alla lettera m) saranno assorbenti dei residui parametri e pertanto destinati a sorreggerli. Così operando otterremo un modello digitale gestito da un algoritmo in grado di dedurre nel colpevole (o presunto che sia) il reale movente posto alla base della sua azione e se tale movente era viziato da qualche patologia genetica o di natura psichiatrica. I dati sui quali i parametri si basano saranno prodotti in funzione delle indagini giudiziarie e dalle perizie e consulenze di natura psicologica e psichiatrica nonché genetica d'attuarsi sul presunto colpevole. In virtù di quanto sino ad ora esposto e più precisamente a quanto si è delineato al punto 3) dei "Materiali e Metodi", anche l'art. 42 c.p. descrittivo del dolo e della colpa dovrà essere rivisto e corretto ove è citata la "coscienza e volontà" dell'azione rettificandola in "autocoscienza" e volontà dell'azione atteso che per essere definiti imputabili occorre la

capacità di “intendere” ovvero di rendersi conto di agire contro l’ordinamento penale (12). Di conseguenza; Art.42 codice penale: “Nessuno può essere punito per un’azione od omissione preveduta dalla legge come reato, se non l’ha commessa con autocoscienza e volontà. Nessuno può essere punito per un fatto preveduto dalla legge come delitto, se non l’ha commesso con dolo, salvo i casi di delitto preterintenzionale o colposo espressamente preveduti dalla legge”.

#### Bibliografia

- (1) Aramini, M. (2002) “I processi inferenziali nel profilo psicologico del criminale” in *Psicologia & Giustizia*, 1. 2-3. 04 Febbraio 2021;
- (2) Turvey B. (1999) “Criminal profiling: An introduction to behavioral evidence analysis”, Accademic Press, San Diego;
- (3) Antolisei F. (2003) “Manuale di Diritto Penale” Giuffrè Editore;
- (4) Lusa V. (2018) ” La valutazione dell’iter criminis nel procedimento penale. La nuova elaborazione dell’elemento soggettivo del reato”, Key Editrice Giuridica, Frosinone;
- (5) Caspi. A. et *All.* (2002) “Role of genotype in the cycle of violence in maltreated children”*in Science*;
- (6) Beach S.R.H. et *All.* (2010) “Child Maltreatment Moderates the Association of *MAOA* with Symptoms of Depression and Antisocial Personality Disorder”*in Journal of Family Psychology* n.24;
- (7) Moore T.M., Scarpa A, et *All.* (2002) “A meta-analysis of serotonin metabolite 5-HIAA and antisocial behavior” in *Aggressive Behavior* n.2;
- (8) Murphy D.L., Lesch K.P. (2008) “Targeting the murine serotonin transporter: insights into human neurobiology”*in Neuroscience*;
- (9) Bamford N.S., Wightman R.M. e Sulzer D. (2018) “Dopamine's Effects on Corticostriatal Synapses during Reward-Based Behaviors”, in *Neuron*, vol. 97, n. 3, 7;
- (10) Lusa V. (2017) ” La valutazione dell’iter criminis e l’esatta configurazione del reato. Un esempio? E’ possibile giuridicamente configurare, o meno, il reato di Femminicidio?”, in *Diritto24 Plus di Guida al Diritto*, e in *Lex24*, Sole24Ore Editore, Milano;
- (11) Lusa V., Franza A. (2019) ”Iter Criminis and Mens rea: A Methodology to understand an Offender’s true intention based on a revision of the concept of guilt in Italian Criminal Law” in *AMERICAN ACADEMY OF FORENSIC SCIENCES*, in *Proceedings in Jurisprudence*, Denver Colorado (USA);
- (12) Lusa V. (2023) “Imputabilità e Pericolosità Sociale alla luce della Riforma Cartabia” Franco Pacini Editore, Pisa.

# Prove tecniche di AI a scuola: come integrarla nel nostro sistema educativo

L'Intelligenza Artificiale sta rivoluzionando l'ambito dell'istruzione, portando a un radicale cambiamento nell'approccio all'insegnamento e all'apprendimento. Il punto su come l'IA si inserisce nel contesto educativo, le implicazioni delle nuove forme di valutazione basate sull'intelligenza artificiale e il ruolo del docente in questo scenario

Di Daniela Di Donato, Docente di italiano (Liceo scientifico), PhD in Psicologia sociale, dello sviluppo e della Ricerca educativa presso Sapienza Università di Roma, esperta di metodologie didattiche, inclusione e uso delle tecnologie digitali a scuola

La scorsa settimana, Google ha lanciato Gemini 1.5, versione potenziata del suo modello di intelligenza artificiale multimodale, cioè in grado di comprendere e operare su diversi tipi di informazioni: testo, codice, audio, immagini e video.

L'Unione Europea intanto, si avvia verso le fasi finali dell'approvazione dell'Ai Act, cioè la prima regolamentazione al mondo sull'intelligenza artificiale.

In questo continuum spazio-temporale, a metà tra Ritorno al Futuro e Matrix, dove si posiziona la scuola italiana?

## Intelligenza artificiale e didattica: un tandem possibile?

L'AI e il suo uso nella progettazione educativa e nella didattica potrebbero trasformarsi in quella forza di gravità, necessaria a riagganciare alla contemporaneità un sistema di istruzione faticosamente in cambiamento e promuovere davvero un autentico avanzamento metodologico e una più rapida transizione digitale?

Da poco esaurito il dibattito sulla didattica per competenze, che in verità è ancora brace accesa, e viva più che mai la misteriosa polemica sulle “scuole senza voto” (che poi il voto ce l'hanno perché la normativa quello prevede), che potrebbero invece diventare presidio di riflessione e avanguardia contro la percepita agonia dell'apprendimento da parte di studentesse e studenti annoiati e, cosa ancora più grave, intorpiditi forse dalla lentezza con cui i cambiamenti riescono a permeare il tessuto educativo, ecco arrivare altre sfide, sempre più complesse.

La verità è che forse non lo sappiamo ancora, dove ci troviamo.

## Gli investimenti nel settore educativo: l'arrivo dell'IA nelle aule



Certamente le scuole hanno ricevuto centinaia di migliaia di euro per l'acquisto di tecnologie, per arredi e ambienti di apprendimento laboratoriali e ora hanno anche le istruzioni operative per disporre dei fondi finalizzati a finanziare centinaia di ore di formazione (DM 66/2023).

Quello che si può osservare è che nonostante sacche di gelosia per una didattica tradizionalmente erogativa è iniziata, da quando se ne ha la disponibilità, una timida ma costante sperimentazione dell'AI generativa conversazionale. Basti dare un'occhiata alle centinaia di esperienze raccontate e condivise in alcuni gruppi Facebook dedicati, primo fra tutti "Intelligenza artificiale a scuola", nato a febbraio del 2023 e con più di duemila docenti animati da un collega, Lorenzo Redaelli, felicemente sospeso su una corda da funambolo tra umanesimo e ingegneria. Lo spazio è dedicato a tutti coloro che vogliono conoscere, applicare e promuovere l'utilizzo dell'Intelligenza Artificiale a Scuola e discuterne le implicazioni a livello etico e sociale.

Lo scopo dichiarato è di portare questa coraggiosa riflessione in classe, affinché gli studenti siano consapevoli dei rischi e delle opportunità di questa tecnologia in evoluzione. Tra i diversi usi che se ne stanno facendo in diversi ambiti disciplinari, troviamo la produzione di immagini da prompt testuali, adattamenti e modifiche di testi, elaborazione di compiti e problemi, fino a generazione di elaborati letterari o visuali in stili diversi, ispirati ad autori di ogni epoca o a correnti artistiche e letterarie.

## Prove e tentativi di lavoro con le AI a scuola

Le prove e i tentativi di lavoro con le AI stanno ancora cercando una sistematizzazione e forse uno dei punti di cerniera che si potrebbe trovare è la connessione tra lo sviluppo dell'apprendimento con l'AI, i processi di insegnamento e apprendimento e la progettazione e costruzione di dispositivi valutativi. La valutazione, nella sua dimensione più sfidante e cioè intesa come esperienza di apprendimento ha infatti bisogno di strumenti, che permettano una personalizzazione, che offrano ancora di esempio, che generino possibilità di supporto e adattività, oltre a offrire orientamento e sostenere lo sviluppo di capacità autoregolatrici.

Gli studi sul feedback e sulla natura dell'intelligenza, sui modelli di impotenza appresa e sullo sviluppo del potenziale di apprendimento, aiutano a sostenere l'idea che nell'area del "non ancora" abile, "non ancora" pronto la fiducia e la capacità di percepirsi efficaci (Dweck, 1998) hanno un ruolo molto significativo.

## Lavorare in classe con l'AI per ripensare ad una didattica inclusiva

Per questi scopi lavorare con AI come ChatGPT 4 e Gemini può diventare un percorso sfidante per ripensare ad una didattica inclusiva, che comporti la trasformazione di alcuni processi cognitivi e un coinvolgimento responsabile delle studentesse e degli studenti.

Se è vero che muoversi nella Zona di sviluppo prossimale con feedback organizzati e costanti di valutazione formativa (Bevilacqua, 2023; Corsini, 2023; Gentile, 2019; Grion, Serbati, & Cecchinato, 2022; Pastore, 2019) permette agli insegnanti di raggiungere tutti gli studenti, le AI possono moltiplicare la possibilità di arrivare a ciascuno di loro, con osservazioni puntuali su compiti, elaborati, risposte, ricerche (Maglioni, 2023). La dimensione pedagogica del corpo che impara insieme alla mente, della multisensorialità talvolta inibita in ambienti di apprendimento poveri di stimoli trova nell'utilizzo dell'AI e delle realtà immersive, ad esempio, un'altra opportunità di restituire la complessità della disabilità e riappropriarsi della autopercezione di sé e delle proprie potenzialità come persone che imparano (De Luca, Domenici & Spadafora, 2023).

## Creare una Rubric di Valutazione con l'IA

Supponiamo di dover creare una Rubric di valutazione per un compito come la recensione di un libro letto. Non basta chiedere alle AI di produrre una Rubric perché, così come ci descrive bene il framework DigComEdu, le competenze più avanzate si sviluppano partendo da abilità già consolidate, da una pedagogia dell'educazione valida e grazie all'aiuto e la collaborazione dei colleghi, con i quali confrontarsi (Di Donato & De Santis, 2021).

Partire da un modello anche semplice, le cui radici siano fondate pedagogicamente, può rendere il processo più accessibile e manifestare meglio i limiti dell'AI e le sue possibilità.

### ChatGPT 4 ▾



**You**

realizza una rubric con almeno quattro criteri per valutare la recensione di un libro destinata a studenti di 13 anni



**ChatGPT**

Creare una rubrica di valutazione per una recensione di un libro destinata a studenti di 13 anni è un ottimo modo per guidare gli studenti nella scrittura di recensioni efficaci e dettagliate. Ecco una rubrica suggerita con quattro criteri principali:

#### 1. **Contenuto e Comprensione:**

- **Eccellente:** La recensione fornisce un riassunto chiaro e completo del libro, mettendo in evidenza i temi principali e i punti chiave senza rivelare spoiler importanti. Mostra una profonda comprensione del materiale.
- **Buono:** La recensione copre la maggior parte dei punti salienti del libro, ma potrebbe mancare di dettagli o profondità in alcune parti. Dimostra una buona comprensione del contenuto.

*Fonte: ChatGPT4, OpenAI*

La conversazione con le AI ha bisogno di praticare la ridondanza: deve cioè svilupparsi un dialogo, attraverso il quale i differenti punti di vista tendano a modificarsi reciprocamente e prendere direzioni sempre più precise. La domanda, il *Prompt*, di chi interroga l'AI diventa più orientato, specifico e l'AI risponde con maggiore proprietà, modellandosi sempre meglio sull'interlocutore e sui suoi desiderata. Il processo però è reciproco, non unilaterale: entrambe le voci, umana e non umana, hanno bisogno di dialogare.

La sintassi del *Prompt* è un linguaggio, la prima grande tecnologia comunicativa dell'uomo e della donna, e paradossalmente per ora quel linguaggio, che ci rende umani, serve per parlare con una entità che definiamo "non umana". Quella cosiddetta *capacità di spostamento*, che ha solo l'umanità, cioè poter parlare di cose che sono remote nello spazio e nel tempo e produrre uno spostamento di ciò che non può essere mostrato (Benanti, 2021) dialoga con qualcosa/qualcuno, che è invece ovunque (nel web) e in nessun luogo specifico.

Manca nello scambio quel processo di sintonizzazione e rispecchiamento emotivo, che ci permette di esprimere o vedere l'empatia durante una conversazione anche mediata (pensate alle videoconferenze durante il Lockdown, a schermi spenti) e di migliorare decisioni e risorse cognitive (Riva & Mancini, 2023), ma una relazione interpersonale sembra invece possibile (Di Donato, 2023). Dobbiamo quindi tenere conto di un interlocutore, che a sua volta si orienta grazie a ciò che raccontiamo, chiediamo, spieghiamo praticando una comunicazione mediata che richiede attenzione e ascolto.

L'ambiguità nell'esempio di aggettivi come "chiaro" o "completo" richiedono di essere sciolti e si hanno almeno due strade: spaccettarli per tirar fuori definizioni più intelleggibili e prive di ambiguità oppure offrire degli esempi.

Specifica meglio ciascun criterio e aggiungi degli esempi



### ChatGPT


Certamente! Ecco una versione più dettagliata dei criteri per la rubrica di valutazione di una recensione di un libro destinata a studenti di 13 anni, con l'aggiunta di esempi per ogni criterio:

#### 1. **Contenuto e Comprensione:**

- **Eccellente:** La recensione presenta un riassunto completo senza spoiler, evidenziando i temi principali, personaggi e l'ambientazione. Ad esempio, "Il libro esplora il tema dell'amicizia attraverso le avventure di due ragazzi in una città fantascientifica".
- **Buono:** La recensione copre i punti salienti, ma manca di alcuni dettagli o profondità. Ad esempio, "Il libro parla dell'amicizia tra due ragazzi, ma non entra molto nei dettagli delle loro avventure".

*Fonte: ChatGPT4, OpenAI*

Naturalmente un'altra abilità è allora ampliare la consultazione e il confronto a distanza tra le due AI disponibili; il procedimento può svelare punti di contatto, ma anche visioni divergenti.

Contenuto e Comprensione	
Eccellente	La recensione presenta un riassunto completo e accurato del libro, evidenziando i temi principali, i personaggi, l'ambientazione e la trama. Il riassunto è scritto in modo chiaro e conciso, senza spoiler.
 <b>Esporta in Fogli</b>	

**| Ad esempio:**

- "Il libro esplora il tema dell'amicizia attraverso le avventure di due ragazzi in una città fantascientifica. I personaggi sono ben sviluppati e le loro relazioni sono complesse e coinvolgenti. La trama è avvincente e piena di colpi di scena."

Fonte: Google Gemini (ex Bard)

## La competenza del docente nell'era dell'Intelligenza Artificiale

La competenza dell'insegnante qui è indispensabile per selezionare, collegare, sintetizzare quello che emerge dal dialogo e dal confronto: se non si conosce bene il funzionamento e lo scopo di una Rubric valutativa e l'obiettivo che si vuole raggiungere, si ignora i destinatari (il loro linguaggio, la loro competenza) risulta quasi impossibile migrare nella fase successiva, ancora *work in progress*.

**Rubric di valutazione per la recensione di un libro letto**  
(Autori: ChatGPT4, Bard, studentesse e studenti, IU Dandd)  
Work in progress

+	Criterio	Esperta	Competente	Apprendista	In corso di acquisizione
	Contenuto e Comprensione	La recensione <sup>1</sup> presenta un riassunto completo e accurato del libro, evidenziando i temi principali, i personaggi, l'ambientazione e la trama. Il riassunto è scritto in modo chiaro e conciso, senza spoiler.	La recensione <sup>2</sup> descrive i punti salienti del libro, ma potrebbe omettere alcuni dettagli importanti o presentare delle imprecisioni. Il riassunto è chiaro (vedi caratteristiche del riassunto) e comprensibile, ma potrebbe essere migliorato in termini di precisione e completezza.	La recensione <sup>3</sup> è basilare e potrebbe omettere dettagli importanti o presentare delle imprecisioni. Il riassunto è difficile da seguire e potrebbe contenere spoiler.	La recensione <sup>4</sup> è troppo generica o sbaglia a raccontare la storia.
	Analisi e Interpretazione	La recensione <sup>5</sup> offre riflessioni personali	La recensione presenta alcune riflessioni personali,	La recensione <sup>7</sup> include poche analisi personali,	La recensione <sup>8</sup> non offre alcuna

<sup>1</sup> Es. "Il libro esplora il tema dell'amicizia attraverso le avventure di due ragazzi in una città fantascientifica. I personaggi sono ben sviluppati e le loro relazioni sono complesse e coinvolgenti. La trama è avvincente e piena di colpi di scena"

<sup>2</sup> Es. "Il libro parla dell'amicizia tra due ragazzi che vivono in una piccola città. I protagonisti sono molto diversi tra loro, ma imparano a conoscersi e a rispettarci. La storia è piena di avventure e colpi di scena. Mi è piaciuta molto, anche se a volte ho trovato la trama un po' lenta."

<sup>3</sup> Es. "Il libro parla di due ragazzi che fanno cose insieme. È una storia semplice, ma è piacevole da leggere. I personaggi sono simpatici e la trama è abbastanza avvincente."

<sup>4</sup> Es. "Il libro parla di due ragazzi. È una storia noiosa."

<sup>5</sup> Es. "Il modo in cui il libro affronta la perdita mi ha fatto pensare alla mia esperienza di quando ho perso un amico".

<sup>7</sup> Es. "Il libro tratta l'amicizia, ma non l'ho trovato molto interessante"

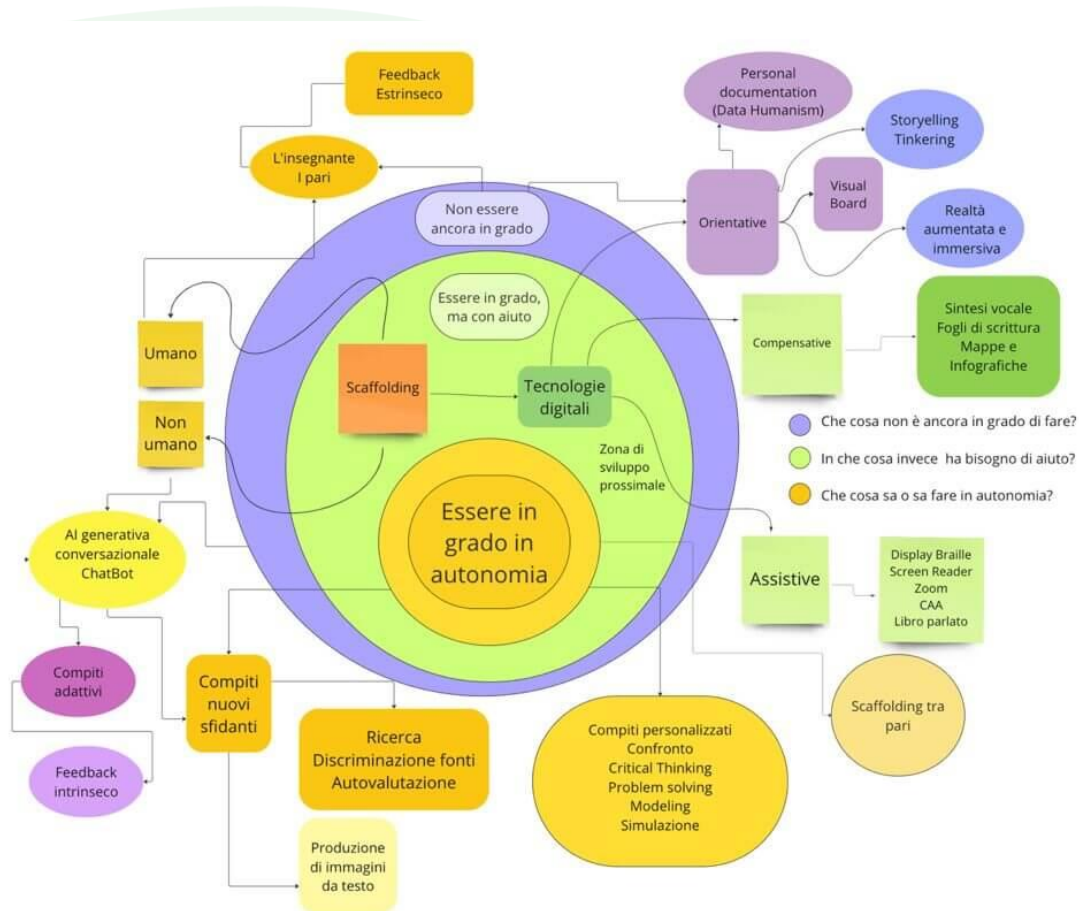
<sup>8</sup> Es. "È solo un libro su due amici"

Fonte: Elaborazione propria

Per passare da una dimensione solo quantitativa ad una versione della Rubric più raffinata e comprensibile, che parli non solo il linguaggio delle AI, ma anche quello di studentesse e studenti destinatari e protagonisti del dispositivo valutativo, ci vogliono ancora altri passaggi e consultazioni, che qui non mostrerò.

## AI a scuola, una costruzione di senso proiettata verso il futuro

Che cosa ci potrebbe suggerire allora questa procedura? In questa necessaria costruzione di senso proiettata verso il futuro, forse occorre prudentemente tornare indietro di qualche passo, su una strada che la scuola dovrebbe conoscere bene e praticare meglio: le questioni metodologiche, le architetture didattiche, le strategie educative e la progettazione dei percorsi di apprendimento per ricominciare a riflettere su quale uso della Zona di sviluppo prossimale vogliamo fare come insegnanti e come sia possibile che le tecnologie digitali diventino leva di sviluppo per inclusione e padronanza.



Sviluppo dell'AI-pprendimento - Fonte: Elaborazione propria

## Conclusioni

Potrebbe essere questo il punto in cui ci troviamo ora? Difficile dirlo con sicurezza: ci muoviamo talvolta un po' confusamente tra le innovazioni, che vorrebbero cambiare la scuola. Intanto però vi lascio la bozza di un modello che vorrebbe essere una lente per osservare i processi (giuro che non l'ha realizzato l'AI) e potrebbe forse rappresentare una fase intermedia di riflessione:

sull'Intelligenza artificiale, sulle tecnologie digitali, sullo sviluppo di dinamiche di relazione per migliorare e integrare apprendimenti, sulle metodologie e le sfide che aspettano noi al varco spaziotemporale, che ancora ci separa dalla scuola che vorremmo abitare.

## Bibliografia

Benanti, P. (2021). *La grande invenzione. Il linguaggio come tecnologia dalle pitture rupestri al GPT-3*. Edizioni San Paolo.

Bevilacqua, A. (2023). *Il feedback a scuola. Strategie per promuovere l'apprendimento*. Carocci editore.

Corsini, C. (2023). *La valutazione che educa. Liberare insegnamento e apprendimento dalla tirannia del voto*. FrancoAngeli.

Dweck, C. S. (2000). *Teorie del sé. Intelligenza, motivazione personalità e sviluppo*. Erickson.

De Luca, C., Domenici, G., & Spadafora, G. (2023). *Per una inclusione sostenibile. La prospettiva di un nuovo paradigma educativo*. Anicia.

Di Donato, D. (2023). You, me and the digital. The triangle of technology-mediated relationships after the Covid-19 pandemic. *Qwerty-Open and Interdisciplinary Journal of Technology, Culture and Education*, 18 (2), 5-13.

Di Donato, D. & De Santis, C. (2021). Il cambiamento delle pratiche didattiche dei docenti italiani durante il lockdown. Percezioni dell'efficacia nell'uso delle tecnologie didattiche digitali e collaborazione con i colleghi. *Ricerca-azione*, 13(1), 213-233.

Gentile, M. (2019). *Valutare per apprendere. Attività e strumenti per il lavoro in classe*. Utet Università.

Grion, V, Serbati, A. & Cecchinato, G. (2022). *Dal voto alla valutazione per l'apprendimento. Strumenti e tecnologie per la scuola secondaria*. Carocci editore.

Maglioni, M. (2023). Due nuovi strumenti per l'inclusione e per la didattica laboratoriale: Bard e ChatGPT. In *La Ricerca. Uomini e Bot. Intelligenti, utili, dannose? Il punto sulle IA*. 11 (25), 71-73.

Pastore, S. (2019). *Autovalutazione. Promuovere la riflessione e l'autoregolazione dell'apprendimento*. Utet Università.

Riva, G., & Mancini, T. (2023). *Psicologia dei media digitali*. Il Mulino.

# Terrorismo islamico: spazialità reali e virtuali

Il terrorismo islamico, attraverso l'uso strategico dei media, ridefinisce le geografie tradizionali, creando spazialità digitali e fisiche che amplificano la sua portata. Attraverso la simbiosi tra terrorismo e comunicazione, la narrazione terroristica influenza così la percezione della sicurezza e destabilizza gli equilibri geopolitici

Di Marino D'Amore, Docente di Sociologia generale presso Università degli Studi Niccolò Cusano

Il terrorismo rappresenta un fenomeno proteiforme e imprevedibile, il cui scopo primo è quello di destabilizzare le istituzioni, neutralizzare le resistenze del potere costituito, depauperando, di fatto, la percezione della sicurezza di un corpus comunitario, sia secondo un livello nazionale sia internazionale. Esso rappresenta una realtà che accompagna storicamente l'esistenza umana, riuscendo spesso a rinnovarsi nelle sue forme e nelle sue azioni. Un concetto che rifugge qualsiasi definizione totalizzante, in quanto privo di un consenso condiviso sulla nozione stessa che ne definisce la natura essenziale e intrinseca.

Il terrorismo è una realtà resa visibile, come detto, anche da quella comunicazione che ne costituisce il racconto e ne ha caratterizzato i presupposti secondo un mero meccanicismo. Una comunicazione che, di fatto, lo globalizza, neutralizzando, come detto, i confini, fisici, politici e culturali, ma soprattutto ridefinisce e costruisce comunità, luoghi e territori nella realtà e nell'immaginario popolare.

In base a un meccanismo comunicativo, ogni nuova minaccia assurge alla ribalta nelle vecchie e nuove spazialità mediatiche, divenendo parte del nostro vissuto quotidiano e alimentando le nostre paure. Infatti, i media, nel garantire l'informazione grazie alla possibilità di sfruttare le nuove tecnologie, sono invasivi e in grado di raggiungere ogni angolo del pianeta, spettacolarizzando immagini e messaggi di terrore, esacerbando, in questo senso, una visibilità inedita e ubiqua e diventando così i principali vettori di traiettorie, spaziali e al contempo semantiche, che destabilizzano le società.

## Introduzione

Tale articolo si propone di analizzare il fenomeno terroristico, di comprendere come, e secondo quali modalità, esso, e la sua narrazione, modifichino le geografie esistenti, ne attualizzino di nuove, come detto, e sovvertono equilibri geopolitici stabilizzati nel tempo. Geografie fisiche e, al contempo immateriali, che rendono ipervisibile la territorializzazione del terrorismo stesso, inteso come una realtà molto articolata che, assecondando tendenze metonimiche, si compone e viene influenzato da diversi elementi, come l'integrazione, le migrazioni e le omologanti istanze della globalizzazione. Questi elementi modificano, irreversibilmente e costantemente, le sue spazialità, intrinseche ed esteriori, attivando un processo imprevedibile che ne amplifica gli effetti e la dimensione drammatica, palesando una commistione che si manifesta particolarmente in una declinazione precisa del fenomeno: il terrorismo islamico che verrà analizzato attraverso la sua eziologia, la sua struttura e le sue deprecabili azioni.

Esso rappresenta il risultato ultimo e maggiormente esaustivo di un processo che, grazie anche all'ipermediazione<sup>1</sup>, ovvero all'effetto moltiplicativo dei media, realizza e costruisce nuove imprevedibili spazialità.

Tale desiderio di immediatezza si manifesta attraverso il tentativo di offrire al pubblico esperienze profondamente reali in cui il mezzo di comunicazione si cela nell'apparenza di un luogo condiviso, ossia nell'elisione stessa del medium, permettendo a chi comunica di entrare in diretto contatto con l'opinione pubblica a cui si elargiscono nuovi panorami virtuali, scenari ammantati di un sedicente fascino ma che nascondono tutte le loro esecrabili criticità.

Questi sono i criteri sui quali si struttura l'azione comunicativa dei terroristi: instaurare un dialogo immediato con il proprio interlocutore, sfruttando le aree, interattive e multimediali, che connotano i new media. Mediante tale meccanismo, infatti, le grandi organizzazioni del terrore riescono a massimizzare gli effetti delle loro azioni, sia in termini di visibilità sia in termini di allarme diffuso. L'opinione pubblica si trova davanti alla forza devastante di immagini di attentati, di sofferenza, di morte e percepisce quelle realtà come molto più prossime e preoccupanti, esacerbate da tale rappresentazione glorificante e autoreferenziale che corrode la dimensione spaziale.

## Terrorismo e comunicazione

Contenuti di carattere intimidatorio vengono condivisi sulla rete da radicali musulmani che, aggrappandosi ad una distorta visione del vero Islam, istigano alla violenza e alla lotta contro gli infedeli. La mole di contenuti di stampo terroristico che spunta sul web ogni giorno è spaventosa e proprio per questa ragione lo sforzo deve essere sempre collettivo<sup>2</sup>.

Per questo motivo, Facebook, YouTube, Twitter e Microsoft si uniscono per la creazione un database condiviso di tracce digitali uniche che vengono assegnate ai contenuti di carattere violento o collegati ad attività terroristiche condivisi online.

In questo modo se una pubblicazione viene bloccata, ad esempio su Facebook, non potrà essere condivisa successivamente su YouTube o su Twitter.

Google menziona l'uso di tecnologie di machine learning per riuscire ad identificare e segnalare in maniera più veloce i contenuti legati, direttamente o indirettamente, ad attività terroristiche.

Nonostante gli sforzi nello sviluppo di queste tecnologie, la principale sfida riguarda ancora la difficoltà nel distinguere i contenuti: un video con immagini o contenuti associati al terrorismo può essere informativo se presentato dalla BBC o dalla CNN ma può essere un potente catalizzatore di condotte violente o estremiste se pubblicato da un altro utente.

Inoltre, Facebook, Microsoft, Twitter e YouTube hanno annunciato anche la creazione del Global Internet Forum to Counter Terrorism, che avrà lo scopo di liberare, in ottica sinergica e collaborativa, i rispettivi servizi di hosting di contenuto dalla comunicazione di estremisti violenti.

Altro strumento della strategia di lotta al terrorismo sul web occorre menzionare anche che YouTube ha introdotto il *redirect method*, che reindirizza chi cerca dei video su contenuti estremistici verso delle playlist che contrastano e demistificano i contenuti stessi e la propaganda terrorista.

I video presenti in queste playlist hanno l'obiettivo di fornire più elementi che possano contribuire a cambiare l'opinione delle persone che sono a rischio di radicalizzazione.

Tutta la comunicazione dello Stato islamico è stata fondata sulla grande narrazione che i gruppi storici jihadisti, soprattutto nel caso di Al-Qaeda, hanno creato dopo la guerra in Afghanistan contro

---

<sup>1</sup> Il sostantivo ipermediazione è la traduzione del neologismo inglese hypermediacy, introdotto nel dibattito mediologico alla fine degli anni Novanta del secolo scorso da Jay David Bolter e Richard Grusin (*Remediation Understanding new media* 1999). Il concetto indica uno dei due modi con cui le immagini si presentano al fruitore dei mezzi di comunicazione contemporanei. Le immagini tendono alla caratteristica dell'immediatezza e appaiono ipermediate, cioè moltiplicano i segni del loro essere mediate, cercando di riprodurre la ricchezza sensoriale dell'esperienza umana

<sup>2</sup> Altheide D. L., *Terrorism and the Politics of Fear*, Rowman & Littlefield, Washington 2017, pp. 141-173.



i sovietici. Questa narrazione include il furto di simboli islamici collettivi come la bandiera nera, ma soprattutto il messaggio fondamentale che riguarda la salvezza e il senso di appartenenza all'unica comunità dei credenti che consente di guadagnare il paradiso attraverso il martirio.

La propaganda del sedicente Califfato ha solo beneficiato della rivoluzione tecnica del Web 2.0 per la sua diffusione nella Rete<sup>3</sup>. Il processo di fidelizzazione allo Stato Islamico è un'operazione che si avvale di alta tecnologia, gestita da professionisti che se ne occupano a tempo pieno. Infatti, quando un contenuto viene rimosso da un sito o da una piattaforma online, esso ricompare subito da un'altra parte, permettendone la fruizione e il caricamento su altri blog, siti, social network secondo dinamiche piramidali. Allo stesso modo, quando un account importante subisce una sospensione su Twitter, Facebook o YouTube, viene ricreato nel giro di pochi minuti e riprende la sua attività senza particolari difficoltà.

La propaganda comprende il formato tradizionale, cartaceo e quello digitale. Nel primo caso, essa risulta utile per i predicatori che si aggirano per le città, distribuendo opuscoli e piccole guide che informano sui risultati realizzati nel rispetto della legge coranica. Nel secondo, serve per comunicare con gli utenti della Rete, impressionandoli con fotografie, video e messaggi audio che inneggiano al Califfato.

Il suo organo di comunicazione ufficiale più noto, come detto, è Dabiq, una rivista pubblicata online, in formato pdf, facilmente scaricabile e condivisibile, il cui primo numero risale al 5 luglio 2014.

Si tratta di una scelta basata, anche in questo caso, su esigenze concorrenziali e sulla necessità di opporsi a Inspire, la rivista di Al-Qaeda che è, ancora oggi, un punto di riferimento per i jihadisti. Il periodico è redatto primariamente in inglese, ma è possibile reperire anche delle versioni in arabo, russo, francese e tedesco secondo istanze diffuse il più possibile globalizzanti.

I responsabili della propaganda hanno anche dimostrato una grande abilità nel gestire la convergenza mediatica, creando interdipendenze comunicative tra diversi device, usufruendo di ogni tipo di piattaforma comunicativa<sup>4</sup>.

Ad esempio, nel febbraio 2014 è stato annunciato il videogioco *Grand Theft Auto: Salil al-Sawarim*, che riprende il titolo di un noto videogioco e lo collega a un video di propaganda pubblicato qualche tempo prima (Salil al-Sawarim), che ha così ottenuto circa 57.000 visualizzazioni. In quest'ultimo caso, si evidenzia anche la padronanza del concetto di gamification, che trae la portata comunicativa e formativa del gioco a favore della diffusione del messaggio di Daesh.

Ogni zona dello Stato, suddiviso in una ventina di province, *wilayat*, ha un proprio centro di produzione, che si occupa delle notizie locali, aggiornando pienamente il concetto di glocalizzazione<sup>5</sup>. Inoltre, esistono degli enti centrali che si occupano della realizzazione di contenuti grafici, testuali o video che vengono riutilizzati dai centri subordinati, ossia *al-Furqan* (in riferimento alla sura aprente del Corano) e il già citato *al-Hayat Media Center*, fondato nel mese di maggio 2014 dal rapper tedesco ed ex foreign fighter Abu Talha al-Almani, alias Deso Dogg.

I contenuti pubblicati o condivisi via internet spaziano dai video di propaganda che mostrano scene di combattimento, chiamate al jihad o testimonianze di combattenti; passando per i *nashid*, ossia inni che cantano le gesta dei mujahidin; articoli relativi agli eventi locali oppure a dichiarazioni delle autorità straniere; fino agli e-book che indicano i passaggi per costruire degli ordigni esplosivi o descrivono l'ideologia dello Stato, come i libri programmatici della serie *Black Flags*, che spesso ricorrono a informazioni provenienti dai media europei o statunitensi, ricontestualizzate iconograficamente, utilizzando immagini "hollywoodiane" di alta qualità realizzate con il fotomontaggio e altri effetti grafici. Produzioni spesso associate a citazioni di figure autorevoli o a passi coranici, che promuovono documentari del gruppo, alcune realizzate con l'ausilio di droni.

<sup>3</sup> Maggioni M., op. cit., pp. 73-90.

<sup>4</sup> Weiss C., Islamic State launches mobile app for children, "The Long War Journal", 11 maggio 2016.

<sup>5</sup> Bauman Z., Globalizzazione e glocalizzazione, Armando editore, Roma 2005.

I canali principali che vengono usati per viralizzare tale materiale, in modo da neutralizzare la censura dei singoli siti, sono i social network, specialmente Twitter, Facebook e YouTube e siti di condivisione gratuita di file come justpaste.it, archive.org, Diaspora e Live Leak.

Questi strumenti permettono di diffondere facilmente i contenuti e di sfruttare le logiche della diffusione internettiana per raggiungere un pubblico sempre più ampio in tempo reale. Inoltre, al contrario di mezzi di comunicazione unidirezionali come le riviste, i social permettono l'interazione e forniscono l'occasione di entrare direttamente in contatto con reclutatori e simpatizzanti, che si dimostrano sempre molto disponibili a rispondere alle domande degli utenti. In questo modo il messaggio dello Stato Islamico può raggiungere e influenzare, attraverso le proprie tecniche di fidelizzazione e la retorica che lo connota, migliaia di potenziali combattenti<sup>6</sup>.

La tecnologia opera secondo un'ottica multitiming e multiplating, in ogni momento e in ogni luogo, impedendo di rimanere estranei a eventi che avvengono in altre parti del mondo. Tale processo rappresenta un'arma a doppio taglio, poiché se le informazioni che riceviamo sono incomplete o edulcorate, è inevitabile una loro decodifica aberrante<sup>7</sup> e la formazione di un'opinione distorta della situazione. Lo Stato Islamico ha compreso che internet, per i cittadini occidentali, sembra più reale della stessa realtà e sta sfruttando questa debolezza per esteriorizzare visioni terribili, polarizzare il conflitto tra fedeli e infedeli e suscitare dibattiti interni volti a neutralizzare il confronto e il dibattito politico che ne consegue.

Il web viene gestito con molta disinvoltura e maestria, utilizzando strumenti per la comunicazione crittografata, a partire dal già citato Telegram fino a sistemi VNP e Tor, che impediscono di identificare l'indirizzo IP da cui si accede alla Rete e che permettono di aggirarsi nel deep web senza lasciare tracce. In seguito agli attentati di novembre 2015 a Parigi, si è anche sostenuto che IS, acronimo per definire lo Stato Islamico dopo le sconfitte e i territori perduti in Siria e Iraq, usasse le chat private della console PlayStation 4 per sfuggire al monitoraggio delle intelligence internazionali, concentrato su devices più attuali<sup>8</sup>. La caratteristica più importante, e al contempo più allarmante, del piano comunicativo terrorista è il fatto che esso estenda la propria presenza in modo tentacolare su ogni piattaforma, sito e social network, archetipo simbolico della sua espansione nel mondo reale, realizzato attraverso la conquista territoriale.

La radicalizzazione è il risultato della propaganda sopraccitata: l'atto terroristico, la minaccia diretta e indiretta, le scene cruente sono ormai la linfa che alimenta il voyeurismo mediatico, finalizzata a rendere visibile l'esteriorità di una strategia molto più complessa che si prefigge l'obiettivo della radicalizzazione su larga scala.

L'aspetto più d'impatto della comunicazione dello Stato Islamico non è però legato alla narrazione militarizzata, ma all'altra faccia della medaglia, quella che vede i membri dell'organizzazione mentre soccorrono i civili e giocano con i bambini, ossia la componente emozionale che inganna, fidelizza e riabilita, umanizzandola profondamente, quella violenza cieca ed esecrabile, negandone la barbarica recrudescenza. La strategia mediatica di IS è quindi più complessa e sottile di quella di altre organizzazioni, che puntano esclusivamente sul terrore puro per conquistare territori e prostrare le loro comunità, confermando il suo elevato livello di assimilazione di strumenti e tecniche comunicative<sup>9</sup>, nonché la sua abilità di adoperarli a danno dei propri nemici.

## Nuove spazialità e territorializzazione digitale

Grazie alla globalizzazione e alla diffusione di internet, la realtà geografica si divide in reale e virtuale: quella reale, che concerne il conflitto vero e proprio, e quella digitale poiché il

---

<sup>6</sup> Maggioni P., Magri P. (a cura di), op. cit.

<sup>7</sup> Eco U., Sulla televisione. Scritti 1956-2015, a cura di Gianfranco Marrone, La Nave di Teseo, Milano 2018, pp. 54-77.

<sup>8</sup> Serafini M., PlayStation e Isis, cosa sappiamo delle comunicazioni del Califfato, "Il Corriere della Sera", 27 novembre 2015

<sup>9</sup> Napoleoni L., Isis. Lo stato del terrore. L'attacco all'Europa e la nuova strategia del Califfato, Feltrinelli, Milano 2016, pp. 88-119.

reclutamento alla jihad avviene principalmente grazie all'utilizzo dei social network, elidendo di fatto ogni caratterizzazione di tipo spaziale e temporale<sup>10</sup>.

La guerra lontana almeno geograficamente appare allo stesso tempo vicina e presente, sia attraverso la vetrinizzazione quotidiana degli accadimenti in televisione o in Rete, con notizie e immagini al riguardo, e sia perché numerosi combattenti dello Stato Islamico provengono in gran numero da città europee: i cosiddetti Foreign Fighters.

L'Islam radicalizzando muta costantemente, passando da una realtà improntata ai desueti stati nazione per strutturarsi secondo un modello profondamente globalizzato. Questa nuova forma si caratterizza per la capacità di reclutare un'enorme massa di seguaci potenzialmente illimitata.

la Rete assurge a medium principe che esteriorizza la sua duplice natura: quella di strumento di socializzazione, ma, al contempo, mezzo che permette di fruire contenuti in maniera individuale. Si sta tornando quindi a una fase in cui il singolo autore di attentati ottiene una visibilità potenzialmente planetaria grazie alle dinamiche, invasive e piramidali, del web.

Lo stesso McLuhan aveva preconizzato questa stretta connessione tra media e terrorismo, definendo quest'ultimo come un modo di comunicare e affermando che senza comunicazione non vi sarebbe terrorismo<sup>11</sup>.

Gli stessi media, soggetti alle logiche del newsmaking e della commercializzazione della notizia sono sempre più propensi a diffondere e dare visibilità a contenuti di questo tipo, all'interno di un mercato profondamente concorrenziale in cui il discriminante è l'esclusività. Elemento che sottolinea il rapporto reciproco, necessario e speculare tra comunicazione e terrore.

La propaganda dello Stato Islamico è consapevole dell'assuefazione dei fruitori rispetto alla visione di eventi scioccanti e spettacolari e di come, troppo spesso, la verità passi in secondo piano rispetto al sensazionalismo della sua narrazione

I media contemporanei, in questo senso, attualizzano un overload informativo che catalizza interesse e pone in essere una sorta di processo ipnotico che fidelizza lo spettatore e ne costituisce i bisogni comunicativi.

Storicamente le società occidentali hanno tentato, attraverso lo sviluppo sociale, culturale e tecnologico, di mitigare la presenza della morte e concretizzare l'idea di una sostanziale sicurezza sociale, rendendo la percezione del pericolo un elemento esiziale.

Tale deficit stimola i terroristi a distruggere un delicato equilibrio attraverso una sfida, prima simbolica e poi fattuale, all'intero sistema sociale. Questo, secondo Baudrillard, è il cosiddetto spirito del terrorismo<sup>12</sup>.

La morte simbolica, una morte portata quindi all'estremo, fa breccia nella sensibilità dello spettatore e instaura in lui il sentimento della paura e dell'insicurezza; è un eccesso di realtà che il pubblico non riesce ad accettare e ne è quindi sconvolto. Molti contenuti fotografici talmente scioccanti da risultare incredibili iniziano a circolare sul Web anche pochissimo tempo dopo un attentato, come ad esempio le drammatiche immagini diffuse su Facebook da un turista italiano sulle Rambla a Barcellona dopo l'attacco del 17 agosto 2017, le quali mostravano i feriti e i morti a terra dopo il passaggio del furgone guidato dagli attentatori.

Le immagini televisive del crollo delle Torri Gemelle nel 2001 erano così intense e incredibili da sembrare quasi estratte da un film; le fotografie scattate durante una guerra che raffigurano ad esempio delle torture inflitte a dei prigionieri non sono più in grado di rappresentare la realtà poiché, rendendo tutto esplicito e visibile, hanno perso la capacità di comunicare.

Attualmente i media, da standardizzati e di massa, si tramutano in media sociali, ossia capaci di consentire la partecipazione attiva dell'utente.

La Rete è diventata quindi un luogo di scambio prima che di mera comunicazione. Il messaggio, se prima passava dalla televisione allo spettatore, essendo in tal modo unidirezionale, ora non arriva

<sup>10</sup> Khanna P., *Connectography. Le mappe del futuro ordine mondiale*, Fazi, Roma 2016, pp. 156-190.

<sup>11</sup> McLuhan M., *Gli strumenti del comunicare*, Il Saggiatore, Milano 1967, pp. 55-80.

<sup>12</sup> Baudrillard J., Codaluppi V., *La seduzione del simbolico*, Feltrinelli, Milano 2017, pp. 43-69.

più a un soggetto tendenzialmente passivo, ma verso un attore comunicativo alfabetizzato in grado di interagire.

La diffusione della Rete sembrerebbe aver quindi democratizzato la società, permettendo a praticamente chiunque dotato di una connessione internet di agire in prima persona, e consente anche a individui con meno risorse di ottenere importanti risultati per quanto riguarda la comunicazione e la costruzione del consenso.

L'influenza dei media è da porre in relazione con il contesto di ricezione e i fattori personali e sociali che lo accompagnano ed essa può variare in base alle identità, alla cultura, alle gerarchie valoriali e alle conoscenze del pubblico.

In questo senso lo Stato Islamico ha iniziato a promuovere un rapporto di scambio con i propri interlocutori, costruendo man mano un legame di fiducia volto a aumentare il consenso, l'aggregazione e la fedeltà.

Lo Stato Islamico è pienamente consapevole della fondamentale importanza che rivestono i suoi sostenitori per la sua stessa esistenza, la sua linfa vitale e, grazie ad una mirata strategia di propaganda, cerca di aumentarne il numero e la visibilità stessa dei propri messaggi, soprattutto grazie all'intervento della propria audience che può fungere da moltiplicatore diffusivo degli stessi presso altri potenziali pubblici.

Un metodo efficace per diffondere i propri messaggi nel web è, ad esempio, il lancio degli hashtag, una tipologia di etichette digitali utili per circoscrivere l'argomento di una conversazione e a farlo circolare soprattutto nei social network a disposizione di coloro che ne siano interessati.

Dal 2009 Twitter ha reso possibile il cosiddetto *live tweeting*, ossia la possibilità di commentare online in tempo reale un evento utilizzando appunto uno specifico hashtag che vi sia collegato<sup>13</sup>.

Nel 2015, nelle ore successive all'attacco terroristico al giornale satirico Charlie Hebdo in Francia, in rete sono comparsi numerosi hashtag che sostenevano apertamente l'azione dei terroristi, come ad esempio #we\_avenge\_the\_prophet, #lone\_wolves, #parisburns. Twitter, in questo modo, è diventato il luogo digitale della narrazione e della contronarrazione: da un lato hashtag che inneggiavano alla violenza terrorista, dall'altro il popolare #jesuischarlie a sostegno delle vittime e del giornale colpito.

La strategia di propaganda islamica punta all'emozionalità, piuttosto che sulla razionalità dell'utente con cui si relazionano, offrendogli numerosi contenuti sottesi a tale connotazione. La rilevanza dei contenuti è infatti un elemento alla base della social media strategy: alimentare l'interesse e la curiosità voyeuristica del proprio pubblico diventa un elemento imprescindibile nell'economia di un atto comunicativo efficace<sup>14</sup>.

I social network si sono rivelati strumenti efficaci non solo per consolidare la reputazione e la credibilità di un gruppo, grazie al loro approccio sempre più consumer oriented, mirato a determinate fasce di pubblico, selezionate in base ai loro interessi, attraverso un'adeguata targettizzazione che seleziona i contenuti rilevanti per l'utente.

Non basta solamente scegliere un titolo d'effetto o uno stile espositivo particolarmente attraente, ma un contenuto, per essere pervasivo, deve attirare l'attenzione del pubblico, risultare di facile comprensione e essere accettato come un fatto reale<sup>15</sup>.

Di fondamentale importanza sono allora le emozioni coinvolte nel processo di persuasione e le tecniche narrative, oltre a coinvolgere il lettore, rendono anche più comprensibile l'argomento trattato.

---

<sup>13</sup> Peverini P., *Social Guerrilla. Semiotica della comunicazione non convenzionale*, Luiss University Press, Roma 2014, pp.120-140.

<sup>14</sup> Giansante G., *La comunicazione politica online*, Carocci, Roma 2014, pp. 145-175.

<sup>15</sup> Hovland, C., Janis, I., & Kelley, H., *Communication and persuasion*, Yale University Press, New Haven 1953, pp. 81-115.

Viene utilizzata a tale scopo dallo Stato Islamico la tecnica del Digital Storytelling, vale a dire l'insieme delle storie personali destinate a una piattaforma online che ora si è in grado di raccontare a un pubblico notevolmente più vasto grazie alla diffusione dei social network e di internet<sup>16</sup>.

Rilevante è infatti la presenza online di numerosi blog dove solitamente sono contenuti i racconti, per lo più in lingua inglese, di chi ha deciso di raggiungere la Siria per unirsi all'ISIS. Nei blog di questi jihadisti si racconta il perché essi si siano affiliati all'organizzazione, di come la loro vita sia notevolmente cambiata in meglio dopo il trasferimento in Siria e, alcuni di essi, suggeriscono anche come comportarsi con le proprie famiglie al fine di non destare sospetti circa la radicalizzazione l'imminente partenza verso il Califfato.

La forza di tali racconti sta nel coinvolgimento del lettore e nella loro modalità di interazione soggettiva e dedicata, fattori che hanno incoraggiato comportamenti imitativi alla base del reclutamento.

## Conclusioni

Il terrorismo, nella sua caratterizzazione omnicomprensiva, è innanzitutto un concetto politico, e come tale, influenzato da fattori storici, culturali, giuridici ed ideologici.

Una definizione universalmente accettata non esiste, anzi è delegata ai diversi orientamenti degli stati e delle organizzazioni governative secondo contingenze e ragioni di opportunità. Un esempio di tale divergenza è riscontrabile nelle diverse valutazioni che Stati Uniti e Europa pongono in essere: Hezbollah non è un gruppo terroristico per il vecchio continente, mentre lo è per americani e israeliani. Invece i mujahidin che combattevano contro Saddam erano terroristi ma sono stati riabilitati dopo un accordo di cooperazione con le truppe occupanti statunitensi.

Il tentativo di sviluppare una definizione esaustiva che non privilegiasse alcun approccio ma che fosse proiettata a integrare le diverse componenti di una visione multidisciplinare non ha prodotto risultati. Tuttavia, all'interno di questa grande varietà semantica possono essere rintracciati elementi comuni, macrotemi che ne costituiscono una sorta di denominatore comune: la violenza; l'uso della minaccia e la diffusione del terrore nella popolazione. Postulati strategici e funzionali al raggiungimento degli obiettivi preposti attraverso nuove geografie, nuove spazialità digitali e le conseguenti narrazioni. Una caratteristica fondamentale che lo connota, soprattutto nelle sue fasi preparatorie, opponendosi alla visibilità, anelata ed esasperata, delle sue gesta, è la clandestinità, ossia la necessità di rendere segreta l'organizzazione e i suoi membri, i quali, in alcuni casi, continuano a vivere apparentemente una vita normale ma sono, in quanto cosiddette cellule dormienti, sempre pronti ad entrare in azione.

Tale segretezza conduce a un altro elemento fondante, corollario di ogni azione: l'imprevedibilità. Una minaccia sconosciuta e poco prevedibile genera angoscia in una comunità e fiacca la sua percezione della sicurezza.

L'aspetto imitativo è altrettanto importante: le pratiche terroristiche si reiterano anche in luoghi e contesti, politici e sociali, diversi: dirottamenti aerei, bombe, cinture e zaini esplosivi sono diventate drammatiche consuetudini negli ultimi anni.

Il simbolismo, come detto, è fondamentale: i significati espliciti e sottesi degli obiettivi, la metacomunicazione e la rivendicazione identitaria sono aspetti centrali nell'eziologia e nell'evoluzione del fenomeno.

Geografie fisiche, e al contempo immateriali, che hanno arricchito di significato la territorializzazione del terrorismo stesso, inteso come una realtà molto articolata che, assecondando tendenze metonimiche, si compone e viene influenzato da diversi elementi, come l'integrazione, le migrazioni e le omologanti istanze della globalizzazione. Tali elementi hanno mutato,

---

<sup>16</sup> Chung-Yin Yeung J., A Critical Analysis on ISIS Propaganda and Social Media Strategies. University of Salford, Manchester 2015, consultabile su [https://www.researchgate.net/publication/316146537\\_A\\_Critical\\_Analysis\\_on\\_ISIS\\_Propaganda\\_and\\_Social\\_Media\\_Strategies](https://www.researchgate.net/publication/316146537_A_Critical_Analysis_on_ISIS_Propaganda_and_Social_Media_Strategies).

irreversibilmente e costantemente, le sue spazialità, intrinseche ed esteriori, attivando un processo imprevedibile che ne ha amplificata, nel tempo, effetti e dimensione drammatica, palesando commistioni particolarmente evidenti in una declinazione precisa del fenomeno: il terrorismo islamico, esplorato attraverso la sua eziologia, la sua struttura e le sue azioni, come risultato ultimo e maggiormente esaustivo di un processo che, grazie all'effetto moltiplicativo dei media, realizza e costruisce nuove spazialità.

Esiste un legame biunivoco tra comunicazione e terrorismo: le nuove tecnologie diventano, come detto, parte integrante del fenomeno, diventandone testimonianza visiva ma, al contempo, concedendosi come ambienti virtuali su cui concretizzare le proprie ideologie e perseguire i propri scopi. Un connubio che si risolve in una macchina propagandista molto efficace, finalizzata a costruire la percezione spettacolarizzata del fenomeno stesso, a mantenere il consenso e a fidelizzare nuovi adepti, mutando significati geopolitici sedimentati nel tempo.

Tale scenario ha esteriorizzato le tecniche comunicative che abitano l'universo semantico del marketing del terrore e, come tali, influenzano prepotentemente le politiche e le geografie mondiali nonché l'assetto, l'organizzazione e l'evoluzione delle medesime. Un processo che palesa una proficua dialettica, intesa in un'accezione drammatica, tra territorializzazione fisica e digitale. Prima Al-Qaeda e poi l'ISIS hanno concretizzato, nel tempo, gli effetti di tale legame, da un lato catalizzando il passaggio e il dialogo tra vecchi e nuovi media e, dall'altro, esacerbando tutte le possibilità comunicative elargite dall'invasività del web, scenario concepito anche da un punto di vista strettamente spaziale come nuovo campo di battaglia su cui combattere le netwars, come insieme di nuove nazioni digitali da colonizzare. Un processo che ha generato nuovi codici, nuovi linguaggi, ma soprattutto nuovi contesti che proiettano un singolo atto in una dimensione globale. Un'immediatezza che si manifesta attraverso il tentativo di offrire al pubblico esperienze profondamente reali in cui il mezzo di comunicazione si cela nell'apparenza di un luogo condiviso, un nuovo panorama virtuale, sfruttando le aree, interattive e multimediali, che connotano i new media.

In questo contesto di interazione globale, che vede neutralizzata la dimensione spazio-temporale, si inserisce il terrorismo internazionale, un fenomeno che si attualizza come uno dei più accaniti ricercatori di promozione nel mondo. A causa della barbarica e prepotente lacerazione della normalità quotidiana di cui si fa portatore, esso si configura, come detto, come un catalizzatore di notiziabilità costante per l'universo mediatico, in tal modo emerge come un fenomeno altamente simbiotico, specialmente nel caso di Al-Qaeda, con la televisione e poi con il web.

La Rete è diventata un luogo di scambio, di interazione e di condivisione, soprattutto per i più giovani, grazie all'utilizzo e alle innovazioni introdotte dai social network. Se queste nuove possibilità di utilizzare la Rete appaiono da un lato rivoluzionarie e innovative, dall'altro hanno permesso la diffusione e la condivisione globale di contenuti deprecabili come quelli che raccontano il terrorismo. Quest'ultimo le utilizza con grande maestria per annientare la percezione della sicurezza, perpetuare la paura nelle società e creare accettazione e consenso alla propria causa. Terrorismo e media, evidenziano un rapporto simbiotico che si traduce in una duplice territorialità: Quella reale dove avvengono gli attentati e le battaglie e quella digitale dove si concretizza la comunicazione e la guerra virtuale. Perciò il fenomeno continua, anche per questo motivo, a rappresentare una grave minaccia, nonostante, nel tempo, abbia perso territori e uomini. La sua presenza in uno spazio fisico può essere combattuta militarmente ma quella internetiana, immateriale, volatile e proteiforme sembra rigenerarsi e moltiplicarsi ogni volta che subisce un attacco e, nei suoi intenti, punta alla conquista del mondo.

#### Bibliografia

- Allam K. F., Il jihadista della porta accanto. Isis, Occidente, Piemme, Roma 2014.
- Amato V., Global 2.0. Geografie della crisi e del mutamento, Aracne, Roma 2012.

- Altheide D. L., *The Media Syndrome*, Left Coast Press Inc, Walnut Creek 2016.
- Altheide D. L., *Terrorism and the Politics of Fear*, Rowman & Littlefield, Washington 2017.
- Avola M., Demirel S., Di Gregorio P., Filii J., Melfa D., Nicolosi G., *A proposito di Charlie. Una riflessione oltre la cronaca*, Euno edizioni, Leonforte 2015
- Baudrillard J., Codaluppi V., *La seduzione del simbolico*, Feltrinelli, Milano 2017.
- Bauman Z., *Globalizzazione e glocalizzazione*, Armando editore, Roma 2005.
- Bauman Z., *Modernità liquida*, Laterza, Bari, 2011.
- Bauman, *Stranieri alle porte*, Laterza, Roma, 2016.
- Cerreti C., Marconi M., Sellari P., *Spazi e poteri. Geografia politica, geografia economica, geopolitica*, Laterza, Bari 2019.
- Chomsky N., Barsamian D., *Sistemi di potere. Conversazioni sulle nuove sfide globali*, Ponte alle Grazie, Milano 2013.
- Giansante G., *La comunicazione politica online*, Carocci, Roma 2014.
- Giddens A., *Il mondo che cambia. Come la globalizzazione ridisegna la nostra vita*, Il Mulino, Bologna 2000.
- Graziano M., *Geopolitica. Orientarsi nel grande disordine internazionale*, Il Mulino, Bologna 2019.
- Guolo R., *L'ultima utopia. Jihadisti europei*, Guerini e associati, Milano 2015.
- Harvey D., *La crisi della modernità*, Il Saggiatore, Milano 2015.
- Hobsbawm Eric J. *La fine della cultura. Saggio su un secolo in crisi d'identità*, Rizzoli, Milano 2014.
- Huntington S.P., *Lo scontro delle civiltà e il nuovo ordine mondiale. Il futuro geopolitico del pianeta*, Garzanti, Milano 2000.
- Hovland, C., Janis, I., & Kelley, H., *Communication and persuasion*, Yale University Press, New Haven 1953, pp. 81-115.
- Khanna P., *Connectography. Le mappe del futuro ordine mondiale*, Fazi, Roma 2016.
- Luizard P., *La trappola Daesh: Lo Stato islamico e la Storia che ritorna*, Rosenberg & Sellier, Torino 2015.
- Maggioni M., Magri P., *Il marketing del terrore. Twitter e Jihad: la comunicazione dell'Isis*, Mondadori, Milano 2016.
- Maggioni M., *Terrore mediatico*, Editori Laterza, Bari 2015.
- Marrone G. (a cura di), *Sulla televisione. Scritti 1956-2015 di Eco U.*, La Nave di Teseo, Milano 2018.
- McLuhan M., *Gli strumenti del comunicare*, Il Saggiatore, Milano 1967.
- Napoleoni L., *Isis. Lo stato del terrore. L'attacco all'Europa e la nuova strategia del Califfato*, Feltrinelli, Milano 2016.
- Pagnotta F. (a cura di), *Linguaggi in rete: conoscere, comunicare nella Web Society*, Le Monnier Università, Firenze 2015.
- Pape R., *Dying to Win: The Strategic Logic of Suicide Terrorism*, New York, Random House, 2005.
- Peverini P., *Social Guerrilla. Semiotica della comunicazione non convenzionale*, Luiss University Press, Roma 2014.
- Picard G. R., *Media Portrayals of Terrorism - Functions and Meaning of News Coverage*, Iowa State University Press, Iowa 1993.
- Picard R. G., *Press Relations of Terrorist Organizations*, in rivista "Public Relations Review", n. 4 1986, pp. 12-23.
- Plebani A., Diez M. (a cura di), *La galassia fondamentalista, tra Jihad armato e partecipazione politica*, Marsilio, Venezia 2015.
- Plebani A., *Jihad e terrorismo*, Mondadori, Milano 2016.
- Popper K. R., *Cattiva maestra televisione*, Marsilio, Venezia 2002.

- Quagliariello G., Sperzi A., Sfida all'Occidente. Il terrorismo islamico e le sue conseguenze, Rubbettino, Catanzaro 2017.
- Roy O., Generazione ISIS. Chi sono i giovani che scelgono il Califfato e perché combattono l'Occidente, Feltrinelli, Milano 2017.
- S. Elden, Terror and Territory, University of Minnesota Press, Minneapolis 2009.
- Schuster M. A., Stein B. D., Jaycox L. H., Collins R. L., Marshall G. N., Elliot M.N., A National Survey of Stress Reactions After the September 11, 2001 Terrorist Attacks, in rivista "The New England Journal Medicine", n. 345, 2001, pp. 1-6.
- Seib P., As Terrorism Evolves: Media, Religion, and Governance, Cambridge University Press, Cambridge 2017.
- Serafini M., L'ombra del nemico. Una storia del terrorismo islamista, Solferino, Milano 2020.
- Soldani F., Geopolitica planetaria dell'impero. I dieci pilastri del capitale finanziario e dell'imperialismo USA, Zambon editore, Jesolo 2017.
- Stone M., Snake and Stranger: Media Coverage of Muslims and Refugee Policy, 2017.
- Talia I., Amato V., Scenari e mutamenti geopolitici. Competizione ed egemonia nei grandi spazi, Patron, Bologna 2019.
- Teti A., Cyber espionage e cyber counterintelligence. Spionaggio e controspionaggio cibernetico, Rubbettino, Cosenza 2015.

#### Sitografia

- Beckett C., Fanning The Flames: Reporting Terror in a Networked World, Tow Center for Digital Journalism della Columbia Graduate School for Journalism 2016, consultabile su [https://www.cjr.org/tow\\_center\\_reports/coverage\\_terrorism\\_social\\_media.php](https://www.cjr.org/tow_center_reports/coverage_terrorism_social_media.php).
- Cannavicci M., Psicologia del terrorismo, articolo, CEPIC- Centro Europeo di Psicologia Investigazione e Criminologia, consultabile su <http://www.cepic-psicologia.it/articolicontributi/#6>.
- Chung-Yin Yeung J., A Critical Analysis on ISIS Propaganda and Social Media Strategies. University of Salford, Manchester 2015, consultabile su [https://www.researchgate.net/publication/316146537\\_A\\_Critical\\_Analysis\\_on\\_ISIS\\_Propaganda\\_and\\_Social\\_Media\\_Strategies](https://www.researchgate.net/publication/316146537_A_Critical_Analysis_on_ISIS_Propaganda_and_Social_Media_Strategies).
- <https://www.hackmageddon.com/2011/06/16/consumerization-of-warfare/> (sito consultato in data 22/05/2020).
- Zanasi A., Gli Internet Centers e le battaglie di Intelligence, Gnosis 2008, consultabile su <http://gnosis.aisi.gov.it/Gnosis/Rivista15.nsf/ServNavig/17> (sito consultato in data 15/08/2020).





# Guerre di algoritmi: sistemi d'arma autonomi e agentività digitale

Il crescente impiego dei sistemi d'arma autonomi negli attuali conflitti sta ridefinendo le dinamiche della guerra moderna, sollevando problemi etici e di accountability. Questo articolo, illustrando la tipologia delle armi dispiegate, esplora fino a che punto tali sistemi possano prendere decisioni indipendenti dal controllo umano, introducendo il concetto di agentività digitale in ambito militare e delineando i pericoli dell'avvento di una digicrazia militare.

Di Giuseppe Galetta, funzionario presso l'Università degli Studi di Napoli Federico II

Nel corso degli ultimi decenni, l'avanzamento tecnologico ha portato a un crescente utilizzo dei sistemi d'arma autonomi nei teatri di guerra. Con l'attuale conflitto in Ucraina l'impiego di tali dispositivi è diventato parte integrante dei piani strategici e delle azioni tattiche sul campo di battaglia, mettendo in evidenza i vantaggi, ma anche i rischi legati ad un loro utilizzo indiscriminato (Meaker, 2023). Questi sistemi d'arma, infatti, grazie all'implementazione di algoritmi di intelligenza artificiale, sono ormai in grado di prendere decisioni autonome e condurre operazioni militari senza un controllo umano diretto (Smith, 2020). La crescente autonomia di tali dispositivi sta rivoluzionando la logica stessa dei conflitti moderni, sollevando interrogativi etici e legali riguardo al controllo e alla responsabilità delle azioni compiute da queste armi intelligenti: è giusto affidare a una macchina la decisione di uccidere? Fino a che punto un sistema d'arma autonomo è in grado di prendere decisioni? E chi è responsabile per tali decisioni?

## Classificazione dei sistemi d'arma autonomi

Gli AWS (*Autonomous Weapon Systems*) o LAWS (*Lethal Autonomous Weapon Systems*) rappresentano una categoria di armamenti progettati per operare in modo autonomo, senza il coinvolgimento diretto di operatori umani nel processo decisionale riguardante l'individuazione e all'attacco dei target. Questi sistemi d'arma utilizzano l'intelligenza artificiale e l'automazione avanzata per eseguire le loro funzioni e possono essere classificati in diverse tipologie, in base alle loro capacità operative e agli scenari in cui sono impiegati:

### Sistemi Droni

Droni armati o UAV (*Unmanned Aerial Vehicles*), ovvero aerei senza equipaggio umano, controllati a distanza o autonomi, utilizzati per svolgere missioni di ricognizione, sorveglianza e attacco, dotati di capacità decisionali autonome, in grado di selezionare e attaccare bersagli senza l'intervento diretto di un operatore umano. I droni di ultima generazione incorporano sistemi FPV (*first person view*): il pilota è in grado di vedere e seguire il bersaglio tramite un visore.

### Sistemi Terrestri Autonomi

Veicoli terrestri autonomi progettati per operare su terreno, dotati di armamenti letali. Possono essere utilizzati per missioni di sorveglianza, pattugliamento o combattimento senza la necessità di un equipaggio umano a bordo.

#### Sistemi Navali Autonomi

Piccole o grandi imbarcazioni con capacità letali senza equipaggio umano, che possono essere impiegate per missioni di pattugliamento marittimo, protezione delle rotte navali, ricerca, attacco o altri scopi navali. I droni marini sono stati utilizzati con successo dall'Ucraina contro la flotta russa nel Mar Nero.

#### Sistemi di Difesa Aerea Autonoma

Sistemi autonomi progettati per identificare e neutralizzare minacce aeree senza l'intervento umano diretto, dotati di capacità avanzate di riconoscimento e decisione basate su algoritmi (*search and destroy*).

#### Robot Terrestri Autonomi

Robot con capacità di movimento su terreno, utilizzati in contesti militari per svolgere varie operazioni, da ricognizioni, a operazioni di sminamento e disinnesco di ordigni esplosivi, a missioni offensive. Tra le più interessanti applicazioni operative di tale tipologia emergono i robot quadrupedi *BigDog*, progettati dall'azienda Boston Dynamics per l'esercito U.S.A. con il finanziamento del DARPA (*Defence Advanced Research Projects Agency*), agenzia governativa del Dipartimento di Difesa degli Stati Uniti incaricata dello sviluppo di nuove tecnologie per uso militare.

#### Sistemi Sottomarini Autonomi

Veicoli sottomarini senza equipaggio umano a bordo o AUV (*autonomous underwater vehicles*), progettati per eseguire missioni sottomarine senza l'intervento diretto dell'uomo.

#### Swarm di Robot

Sistemi di Sciame (*Swarm Robotics*), ossia gruppi di robot o droni che operano in modo coordinato, spesso utilizzati per compiti distribuiti e collaborativi.

#### Armamenti Autonomi

*Lethal Autonomous Weapon Systems (LAWS)*, ovvero sistemi dotati di armi letali, che possono prendere decisioni autonome riguardo l'identificazione e l'attacco di obiettivi, utilizzando sistemi biometrici di riconoscimento.

#### Sistemi di Supporto Logistico

Sistemi logistici autonomi, ossia veicoli o droni progettati per il trasporto di materiali, munizioni, medicinali e altre risorse logistiche senza il coinvolgimento diretto umano.

## L'evoluzione degli AWS/LAWS

Negli ultimi anni sono state proposte varie classificazioni dei sistemi d'arma autonomi, basate sulle caratteristiche tecnico-operative di tali dispositivi e loro sulle prestazioni sul campo di battaglia (Hagström, 2016), ma è importante notare che il panorama degli AWS/LAWS è in continua evoluzione e nuove categorie e sottocategorie stanno emergendo in risposta agli avanzamenti tecnologici e ai mutevoli scenari operativi, anche in considerazione degli sviluppi dell'intelligenza artificiale. Ad esempio, si sta attualmente lavorando alla miniaturizzazione dei dispositivi per renderli difficilmente individuabili o tracciabili, utilizzandoli in operazioni di ricognizione ed

intelligence oltre le linee nemiche (come nel caso dei nano-droni di tipo stealth appartenenti alla classe *Black Hornet*, addestrati ad operare secondo la logica di sciame), cercando al tempo stesso di aumentare la loro capacità distruttiva, come nel caso delle *loitering munition* o munizioni circuitanti (note come droni-kamikaze), che stanno avendo grande impatto sui campi di battaglia in Ucraina (Bode & Watts, 2023). Oggi il campo di battaglia si configura come un orizzonte multidominio, dove si combatte contemporaneamente tra terra, mare, aria, spazio e cyberspazio, e le *smart weapons*, grazie alla loro flessibilità, rapidità d'azione e precisione, rappresentano una risorsa decisiva e irrinunciabile per i paesi belligeranti (Gregory, 2011; Sabry, 2021; Leonard, 2023).

## Le caratteristiche principali dei sistemi d'arma autonomi

Le caratteristiche principali dei sistemi d'arma autonomi possono essere così sintetizzate:

### Capacità decisionale autonoma

Gli AWS/LAWS sono progettati per prendere decisioni autonome e portare a termine azioni letali senza intervento umano diretto, basandosi su algoritmi avanzati e reti neurali artificiali. Questa autonomia decisionale può coprire l'identificazione di obiettivi, la selezione dei bersagli e l'esecuzione degli attacchi, implementando sistemi avanzati di analisi dei dati biometrici degli obiettivi.

### Intelligenza Artificiale avanzata

Gli algoritmi di intelligenza artificiale, inclusi quelli basati su apprendimento automatico e reti neurali, vengono utilizzati per consentire agli AWS/LAWS di apprendere dalle situazioni o da altri sistemi simili in tempo reale, connettendosi anche a risorse OSINT (*Open Source Intelligence*), al fine di adattarsi alle mutevoli condizioni del campo di battaglia (Layton, 2021; Dresch-Langley, 2023). Questa tecnologia permette una maggiore flessibilità operativa e una risposta più adatta alle varie situazioni di combattimento.

### Capacità di integrazione

Gli AWS/LAWS possono essere integrati in reti complesse di comunicazione e sorveglianza, consentendo una coordinazione efficace con altri sistemi d'arma e risorse militari. Questa logica di "sciame" (*swarming*) consente un coordinamento più efficace sul campo di battaglia, permettendo a tali sistemi di agire all'unisono.

### Sensori avanzati

Gli AWS/LAWS sono dotati di sensori avanzati, come telecamere, termocamere, radar, lidar, infrarosso e altro ancora, che consentono di percepire l'ambiente circostante e individuare i bersagli con estrema precisione.

### Velocità di decisione

Gli algoritmi utilizzati dagli AWS/LAWS e la rapidità nell'elaborazione dei dati permettono a tali sistemi d'arma di prendere automaticamente decisioni ad elevatissima velocità, superando la capacità umana di reazione.

Date le elevate capacità di autonomia di tali sistemi d'arma, il dibattito sull'utilizzo degli AWS/LAWS è molto complesso, coinvolgendo aspetti etici e legali, oltre che strategico-politici. Organizzazioni internazionali, ricercatori e attivisti stanno cercando di stabilire principi e regolamenti in grado di mitigare i rischi associati a questa nuova generazione di armamenti, con l'obiettivo di garantire che tali dispositivi possano essere impiegati in maniera responsabile e conforme alle norme internazionali (Allenby, 2015; Williams & Scharre, 2015; Marsili, 2021).

## Contestualizzazione storica e attuale

Il dibattito sull'impiego di sistemi d'arma autonomi non è nuovo. Già nel 1950, lo scrittore di fantascienza Isaac Asimov anticipava le sfide etiche legate all'automazione delle decisioni belliche nella sua opera "I, Robot" (Asimov, 2021), mentre negli anni Sessanta del secolo scorso la telechirica muoveva i suoi primi passi (Clark, 1963; 1964). Tuttavia, solo nelle ultime due decadi, con l'avanzamento dell'intelligenza artificiale e delle tecnologie robotiche, il tema è diventato centrale nei discorsi strategici, politici e militari (De Landa, 1996; Heisbourg, 1999; Payne, 2021). La diffusione dei droni (UAV, *Unmanned Aerial Vehicles*), infatti, ha rappresentato un punto di svolta nella concezione delle operazioni belliche (Kaag & Kreps, 2014; Gusterson, 2016; Ruschi, 2016; Tognocchi, 2018). Il processo di dronizzazione dei conflitti ha raggiunto la sua massima espressione con l'attuale guerra in Ucraina. Tale fenomeno ha imposto una revisione delle strategie militari, ma anche dei principi etici nella conduzione delle operazioni sul campo. Infatti, come evidenziato da Adams (2001), negli ultimi conflitti si sta assistendo ad un progressivo abbandono del processo decisionale umano a fronte di un corrispettivo incremento dell'autonomia dei sistemi d'arma guidati da algoritmi di intelligenza artificiale. Tutto questo solleva interrogativi sulla responsabilità morale e giuridica delle azioni militari, poiché le macchine "intelligenti" stanno assumendo un ruolo sempre più attivo e autonomo nell'uso della forza letale (Chamayou, 2015; Schwarz, 2015; Jaffer, 2016; Johnson, 2019).

## Gli algoritmi e la geografia della guerra

Amoore (2009) ha sottolineato la crescente importanza degli algoritmi nella geografia della guerra contemporanea: l'algoritmo diventa centrale nella selezione dei bersagli e nell'attuazione di azioni letali. Questo solleva preoccupazioni etiche sulle decisioni autonome delle macchine che coinvolgono vite umane. Amoroso e Tamburrini (2019) hanno affrontato la questione dell'etica e del diritto nell'uso crescente dei sistemi d'arma autonomi: secondo gli autori, la necessità di definire il ruolo del controllo umano in tali contesti diventa cruciale per evitare abusi e garantire il rispetto dei principi etici fondamentali, in accordo con le leggi e i trattati internazionali. Panwar (2020) ha approfondito il concetto di autonomia, sottolineando come le armi autonome debbano essere considerate "schiave" piuttosto che "padrone". Questo approccio riflette la necessità di sottoporre le armi autonome ad un controllo umano "significativo" (HRW, 2016). Le stesse istituzioni militari hanno affrontato l'argomento: in Italia, ad esempio, i recenti rapporti del CASD (Centro Alti Studi per la Difesa) e dell'IRAD (Istituto di Ricerca e Analisi della Difesa) illustrano lo stato dell'arte dell'intelligenza artificiale e dell'applicazione dello *swarming* nei contesti militari (CASD-IRAD, 2022a, 2022b). In ambito internazionale, l'International Committee of the Red Cross, riscontrando l'enorme impatto sul campo di battaglia e l'ingente numero di vittime civili coinvolte in attacchi su obiettivi militari (i cosiddetti "danni collaterali"), ha esaminato gli aspetti tecnici, legali e umanitari dei sistemi d'arma autonomi (ICRC, 2014), mentre il SIPRI (*Stockholm International Peace Research Institute*) ha svolto ricerche approfondite sull'evoluzione e il limite dell'autonomia nelle armi (Boulanin & Verbruggen, 2017; Boulanin et al., 2017; 2020).

## Il dilemma etico di Asaro

Il tema dell'*accountability* è il principale nodo da sciogliere quando si parla di sistemi d'arma autonomi (Seixas-Nunes, 2022a). Il filosofo della scienza Peter Asaro si è occupato di questioni legate all'etica e alle implicazioni sociali della robotica e dell'intelligenza artificiale in ambito militare, sollevando un importante dilemma etico alla base dell'impiego dei sistemi d'arma

autonomi nei teatri di guerra (Asaro, 2012). Infatti, la possibilità che le macchine possano prendere decisioni autonome senza il diretto coinvolgimento umano durante le operazioni militari, comporta il rischio di disumanizzazione della guerra, sollevando dubbi circa la responsabilità nella conduzione di azioni letali: può una macchina essere lasciata libera di uccidere in base a un algoritmo? Il nucleo del dilemma è rappresentato proprio dal timore che gli AWS/LAWS, possano essere dotati di un'autonomia decisionale così avanzata da poter selezionare e ingaggiare bersagli senza il controllo umano diretto, diventando di fatto dei *killer robots* (Sparrow, 2007; Armin, 2009; Balistreri, 2017; Felt, 2020): la necessità di preservare i diritti umani e il controllo umano significativo sulle macchine impone dunque la scelta di vietare o meno l'utilizzo dei sistemi d'arma autonomi, identificando limiti internazionalmente condivisi (Krishnan, 2009, Anzarouth, 2021). Tale dilemma solleva dunque molte preoccupazioni etiche, tra cui:

#### Mancanza di responsabilità

Se un sistema d'arma autonomo prende decisioni di vita o di morte senza un controllo umano diretto, chi è responsabile in caso di danni collaterali o errori?

#### Mancanza di etica

Gli algoritmi che guidano tali sistemi potrebbero non essere in grado di comprendere appieno le complessità etiche delle situazioni di combattimento, portando a decisioni che potrebbero essere considerate immorali o controverse.

#### Rischio di abusi

Un'arma autonoma potrebbe essere utilizzata in modo improprio o illegale, ad esempio, per reprimere proteste civili o per fini terroristici.

#### Velocità delle decisioni

Le macchine possono prendere decisioni processando dati ad altissima velocità, il che potrebbe portare a situazioni in cui il tempo per una riflessione etica di tipo umano è limitato o inesistente.

#### Mancanza di empatia

Le macchine non possiedono empatia come gli esseri umani, e questo potrebbe influenzare le decisioni in modo diverso rispetto a un operatore umano, causando morti inutili in determinati scenari operativi.

Peter Asaro e altri esperti in etica dell'intelligenza artificiale sottolineano l'importanza di porre limiti e introdurre normative chiare sull'utilizzo delle armi autonome per ridurre questi rischi etici (Reeves & Johnson, 2014; Stauffer, 2020; Seixas-Nunes, 2022c). Con l'aumento dei conflitti in ogni parte del mondo, le questioni legate alla roboetica, al diritto internazionale e alle leggi sulla guerra sono diventate sempre più rilevanti nel dibattito circa l'impiego di tali tecnologie nelle operazioni militari.

## L'illusione della coscienza artificiale di Arkin

La possibilità di dotare i sistemi d'arma autonomi di una coscienza artificiale in grado di guidare la macchina verso decisioni conformi alle regole internazionali è alla base dell'approccio di Ronald Arkin, che propone un modello di sistema etico di IA di tipo top-down apparentemente in grado di aumentare la sensibilità etica della macchina rispetto alla decisione meccanicamente determinata di uccidere (Arkin, 2009). Secondo Arkin, un sistema d'arma autonomo dovrebbe essere basato su un'architettura IA governata da un "agente morale artificiale" o ATA (Wallach & Allen, 2009) che implementi regole d'ingaggio e vincoli ben precisi, basati su sistemi normativi universalmente

accettati in guerra (come il diritto bellico, la Convenzione di Ginevra e la dichiarazione dei diritti dell'uomo promulgata dall'ONU), in grado di guidare le decisioni della macchina (CASD, 2022c). In base a tale assunto, un sistema d'arma autonomo che incorpori tali regole, dovrebbe essere in grado di verificarle prima di prendere una decisione letale e, quindi, agire in maniera conforme ai principi etici del diritto internazionale umanitario. Purtroppo, al momento attuale non è ancora possibile garantire che la macchina, trovandosi ad operare in specifici contesti che richiedono rapidità di reazione, interpreti correttamente tali regole, anche se universalmente condivise, per cui la soluzione proposta da Arkin appare tuttora di difficile attuazione. Gli studi sullo sviluppo di una coscienza artificiale sono ancora in corso ed altre soluzioni sono state proposte affinché i sistemi d'arma autonomi possano al più presto acquisire una capacità decisionale improntata al rispetto delle regole di guerra, evitando ambiguità interpretative ed impedendo uccisioni indiscriminate o danni collaterali (Strawser, 2013). A tal proposito, un'interessante disamina di vari casi di studio sullo sviluppo di una coscienza artificiale è offerta da Chella (2023).

## L'agentività digitale nei teatri di guerra

L'agentività è un concetto elaborato dallo psicologo Albert Bandura nell'ambito della teoria social-cognitiva e dell'autoefficacia. Si tratta della capacità dell'essere umano di agire attivamente e trasformativamente in determinati contesti operativi, mettendo in atto azioni adeguate allo scopo (Bandura, 2000). L'intenzionalità alla base delle azioni è quindi modulata sulla base delle variabili di contesto, che modificano i livelli di risposta dell'agente, il quale può intervenire in maniera causale sulla realtà per modificarla producendo gli effetti voluti. Se l'agente non è umano ma è una macchina, è chiaro che tutte le possibili variabili dovrebbero essere già inserite nel set algoritmico in base al quale il sistema d'arma autonomo, trovandosi di fronte ad una determinata situazione operativa che imponga delle scelte, possa agire in maniera rapida e adeguata al contesto, conformemente alle istruzioni ricevute.

Il concetto di "agentività digitale" in ambito militare si riferisce alla capacità dei sistemi d'arma autonomi di esercitare una capacità decisionale indipendente rispetto all'operatore umano attraverso l'utilizzo di algoritmi di intelligenza artificiale (Scharre, 2018). L'IA, infatti, svolge un ruolo chiave nell'analisi predittiva e nel decision making durante le operazioni militari: i modelli predittivi basati su algoritmi di apprendimento automatico possono valutare scenari complessi, anticipare movimenti avversari e ottimizzare le risposte alle minacce in tempo reale, utilizzando anche risorse OSINT in tempo reale (Galetta, 2023). Grazie allo sviluppo di capacità predittive e di autoapprendimento, tali sistemi possono assumere un crescente livello di autonomia nella pianificazione, nell'esecuzione e nella valutazione delle azioni senza una supervisione umana diretta in tempo reale, sollevando dubbi cruciali circa la sicurezza, l'etica, la legalità e i profili di responsabilità nella condotta delle operazioni militari (Bergen & Rothenberg, 2014; Wang & Zhang, 2021; Sisinna, 2022). Infatti, se da una parte tali sistemi offrono una maggiore efficacia tattica e operativa sul campo di battaglia, dall'altro la possibilità di errori algoritmici o di compromissione dei sistemi informativi da parte di agenti esterni al sistema (*hacking*) aumentano il rischio di danni collaterali (Chiriatti, 2021).

## Elementi chiave del concetto di agentività digitale in ambito militare

Esplorando l'uso dei sistemi d'arma autonomi nei teatri di guerra, gli elementi chiave della discussione sul concetto di agentività digitale in ambito militare sono i seguenti:

Algoritmi e decisioni autonome

Gli algoritmi sono essenziali per la costruzione di una “agentività digitale”. I sistemi d'arma autonomi utilizzano algoritmi complessi per elaborare informazioni, identificare obiettivi, prendere decisioni ed eseguire azioni senza un intervento umano diretto (Layton, 2018).

#### Apprendimento automatico e adattamento

Alcuni sistemi d'arma autonomi sono in grado di apprendere dai dati e adattarsi alle nuove informazioni senza l'intervento umano. Questo può comportare una maggiore flessibilità operativa, ma solleva anche preoccupazioni riguardo alla trasparenza e alla comprensibilità delle decisioni prese da tali sistemi.

#### Capacità di risposta rapida

L'agentività digitale consente ai sistemi autonomi di reagire rapidamente a situazioni dinamiche e complesse sul campo di battaglia. Questo è spesso considerato un vantaggio strategico, ma può rendere molto difficile evitare decisioni impulsive o errate.

#### Sfide etiche e legalità

L'agentività digitale nei sistemi d'arma solleva importanti questioni etiche e legali. Chi è responsabile per le azioni compiute da un sistema autonomo? Come possono essere garantiti il rispetto del diritto internazionale e delle leggi umanitarie?

#### Controllo umano significativo

In molte discussioni, si sottolinea l'importanza di mantenere un “controllo umano significativo” su tali sistemi. Questo implica che, mentre i sistemi possono avere una certa autonomia, dovrebbero rimanere soggetti al controllo umano in caso di decisioni critiche e per evitare errori e comportamenti imprevisti o indesiderati (Horowitz & Scharre, 2015; Ekelhof, 2019).

#### Interconnessione e “swarming”

In situazioni tatticamente rilevanti, i sistemi d'arma autonomi possono operare in modo coordinato e collaborativo, formando sistemi interconnessi noti come *swarm*. Questi possono amplificare l'agentività digitale di un singolo sistema connettendolo con altre unità sistemiche per produrre un'azione coordinata (logica di sciame), richiedendo quindi una gestione più attenta e controllata (Kallenborn, 2021).

Ad ampliare il concetto di agentività digitale in ambito militare contribuiscono anche gli studi di *swarm robotics*, ovvero l'applicazione della logica di sciame ai sistemi d'arma autonomi, che consentirebbero alle armi intelligenti di agire in armonia durante le operazioni militari, introducendo l'idea di un'agentività collettiva digitale, possibile preludio di una coscienza robotica collettiva (Arquilla & Ronfeldt, 2000). Tale cooperazione tra algoritmi consentirebbe agli eserciti di superare le limitazioni umane nella pianificazione e nell'esecuzione delle operazioni, permettendo di adattarsi automaticamente alle mutevoli condizioni del campo di battaglia ed imponendo una supremazia strategica sul nemico (Singer, 2009a; 2009b). Anche tale prospettiva solleva molti dubbi etici, ma anche preoccupazioni per il futuro stesso dell'umanità (Asaro, 2008; Docherty, 2012; Balistreri, 2015; Alexandre, 2018).

## I pericoli di una digicrazia di guerra

La possibilità che i sistemi d'arma autonomi possano superare i limiti decisionali imposti dall'uomo comporta il rischio che le macchine riescano un giorno assumere il pieno controllo dei conflitti, imponendo una sorta di potere immateriale con conseguenze inimmaginabili, giungendo alla singolarità tecnologica (Kurzweil, 2022). Il primo attacco documentato di un sistema d'arma



effettuato in piena autonomia è stato riportato dalle Nazioni Unite nel 2020 in Libia, nel corso della seconda guerra civile (Hambling, 2021; Vincent, 2021). L'attuale impiego degli AWS/LAWS sul campo di battaglia, in grado di eseguire *targeted killings* in piena autonomia, sta estendendo la portata strategica della guerra oltre i confini convenzionali (Grondin, 2013). Satelliti, droni ed altri dispositivi ad alta tecnologia forniscono una visione dettagliata e in tempo reale del campo di battaglia: questa ubiquità sensoriale consente una comprensione avanzata del teatro operativo, permettendo ai sistemi d'arma autonomi di prendere decisioni rapide e mirate, ma non certo prive di errori e danni collaterali. Si tratta di una digicrazia esercitata attraverso il potere degli algoritmi di IA applicati a strumenti di guerra (Freedman, 2017). Si pensi ad esempio al sistema *Lavender*, che ha guidato i raid israeliani nella striscia di Gaza: si tratta di un software di intelligenza artificiale in grado di selezionare autonomamente gli obiettivi da colpire sulla base di una serie di parametri (tra cui anche il riconoscimento facciale), assegnando un punteggio; coloro che hanno un ranking elevato diventano automaticamente un bersaglio, che viene eliminato dopo aver acquisito l'autorizzazione a procedere da parte di un agente umano. Purtroppo, la lista di obiettivi prodotta automaticamente dal sistema è stata spesso considerata come frutto di decisioni umane anziché uno strumento di supporto alle decisioni, eludendo in tal modo il processo di verifica diretta delle informazioni, con il risultato di causare ingenti danni collaterali e un numero altissimo di vittime innocenti. Il fattore umano è quindi decisivo quando si tratta di eliminare un obiettivo umano, seppur legittimo: come riconosciuto dalle stesse gerarchie militari, il potere decisionale non può essere delegato agli algoritmi, consentendo ai sistemi d'arma autonomi di colpire sulla base di dati che possono rivelarsi fallaci (U.S. Department of Defence, 2023; U.S. Department of State, 2023; Otto & Mănescu, 2023). È quindi necessario rivedere su nuove basi il rapporto tra personale militare e macchine, giungendo ad un nuovo modello relazionale collaborativo in cui l'IA funga da sistema di supporto decisionale piuttosto che da sistema decisionale autonomo (Rosapepe, 2020; Sariel, 2021).

Il termine “digicrazia” utilizzato in ambito militare denota la creazione di un sistema di potere digitale per mezzo del quale le tecnologie avanzate e l'IA giocano un ruolo predominante nella conduzione delle operazioni militari (Arquilla & Ronfeldt, 1997). Tale potere si manifesta attraverso la totalizzazione del controllo del tempo e dello spazio nelle operazioni militari, grazie ad un'estesa rete di sistemi di sorveglianza geospaziale globale (Gregory, 2011; Krieger, 2018), ma anche attraverso la capacità delle armi intelligenti di agire senza dare il tempo all'avversario di accorgersi della minaccia e prendere le opportune contromisure, aumentando la supremazia strategica sul campo. Un drone, infatti, proietta la sua forza senza proiettare vulnerabilità, diventando un dispositivo di potere incarnato (*videre nec videri*): un “Golem digitale” che, ribellatosi al suo creatore, è in grado di esprimere una propria autonoma agentività (Čapek, 2015; Marazzi, 2012; Payne, 2021). Il drone è ormai diventato l'incarnazione dell'avversario sul campo di battaglia: a lui è anche possibile arrendersi, come già fatto da molti soldati russi di fronte ai droni ucraini. Il processo di dronizzazione degli attuali conflitti, in particolar modo quello in Ucraina, sta avendo una portata pervasiva dal punto di vista delle scelte di politica militare degli stati coinvolti, impattando addirittura sui sistemi formativi dei rispettivi paesi, che si sono posti l'obiettivo di educare le giovani generazioni alla prontezza al combattimento: la Russia sta infatti sviluppando un programma di formazione avanzata sui droni, portato avanti dalla Rostec, una holding statale attiva nei settori strategici della difesa e dell'alta tecnologia, che ha lanciato un centro di formazione per addestrare specialisti di UAV; inoltre, è stato reintrodotta nei programmi di studio l'addestramento militare di base per gli studenti delle scuole superiori nei gradi 10 e 11, che comprende nozioni di pilotaggio dei droni da combattimento, allo scopo di dotare i liceali delle competenze necessarie ad operare in scenari di guerra complessi, acquisendo così la qualifica di “combattente di droni”. Anche l'Ucraina sta implementando la formazione sui droni nei programmi scolastici, introducendo corsi di pilotaggio di massa, evidenziando in tal modo il livello della competizione in atto con la

Russia per il dominio della tecnologia UAV, nonché il grado di dipendenza dai droni da parte dei paesi belligeranti nel corso dell'attuale conflitto.

Un ulteriore elemento di supremazia strategica è dato dall'implementazione nei sistemi d'arma di algoritmi predittivi, alimentati da dati in tempo reale, che consentono la previsione di movimenti avversari, il monitoraggio delle risorse logistiche e la valutazione delle condizioni meteorologiche, influenzando l'esito delle operazioni sul campo di battaglia. La digicrazia militare è interamente basata su algoritmi avanzati che conferiscono agentività digitale ai sistemi d'arma autonomi. Questi algoritmi, che sfruttano le capacità di apprendimento automatico dell'IA, consentono a droni, robot e *smart weapons* di poter assumere un ruolo attivo nelle decisioni operative sul campo di battaglia, del tutto svincolate dal controllo umano (Sabry, 2021). L'agentività digitale rappresenta quindi la capacità di esprimere autonomamente intenzioni e azioni da parte della macchina, creando un nuovo concetto di guerra, dove a combattere sono gli algoritmi (Arkin, 2007). In un simile scenario di digicrazia militare, dove ad esercitare il potere sul campo sono i sistemi d'arma autonomi, si assiste ad una trasformazione delle guerre convenzionali in guerre di algoritmi: le forze armate si scontrano non solo attraverso la potenza di fuoco convenzionale, ma anche attraverso la competenza algoritmica. Si tratta di una "guerra surrogata", dove gli eserciti delle grandi potenze militari competono nello sviluppo e nell'impiego di algoritmi avanzati per ottenere vantaggi strategici e operativi, utilizzando anche lo spazio come ultima frontiera di combattimento attraverso il lancio massiccio di satelliti militari, in grado di interfacciarsi con i sistemi d'arma autonomi sulla terra (Krieg & Rickli, 2019).

La digicrazia militare solleva dunque importanti questioni etiche riguardo al controllo umano sulle macchine intelligenti (Sharkey, N.E., 2012; Bode & Watts, 2023). Come già detto, la decisione di affidare a sistemi autonomi la selezione e l'attacco di obiettivi apre un dibattito sulla responsabilità e sulla legalità delle azioni compiute da tali dispositivi. La necessità di regolamentazioni chiare e di linee guida etiche diventa quindi imperativa al fine di mitigare rischi e abusi (Sharkey, A., 2019; Sharkey, N.E., 2016; Marsili, 2021). Attualmente il dibattito internazionale su come debba essere mantenuto il controllo umano sull'uso dei sistemi d'arma autonomi sembra essere regolato da tre comunità di pratica, avente ciascuna i propri obiettivi e priorità: si tratta dei diplomatici, dei produttori di armi e dei giornalisti (Bode, 2024). Tali comunità di pratica sono in grado di modificare l'agenda delle discussioni in atto sull'eticità dell'utilizzo dei sistemi d'arma autonomi e di modellare il dibattito pubblico sulle leggi da adottare, condizionandone l'esito politico. In conclusione, la digicrazia militare rappresenta la nuova frontiera nell'evoluzione della guerra, dove il potere digitale si fonde con le strategie militari e con le lobby d'interesse, influenzando le decisioni dei governi. Se i sistemi d'arma autonomi offrono vantaggi in termini di efficienza e precisione in battaglia, dall'altro richiedono un'analisi multidisciplinare che integri controlli etici, decisioni normative e un dialogo politico internazionale per garantire un utilizzo responsabile di tali tecnologie avanzate (Bousquet, 2009). L'equilibrio tra l'efficienza operativa data dalla tecnologia e la necessità di preservare i valori e i diritti umani fondamentali rappresenta, infatti, una sfida cruciale. La riflessione critica su questo paradigma emergente è essenziale per guidare lo sviluppo futuro delle tecnologie militari su basi eticamente responsabili e internazionalmente condivise, affinché l'obiettivo delle ricerche scientifiche e tecnologiche sui sistemi d'arma autonomi non sia solo quello di accelerare la *kill chain* nel corso delle guerre, diventando una catena di montaggio distruttiva e disumanizzante (Cheater, 2005; Clapper et al., 2007). Il controllo umano su decisioni letali, la prevenzione di abusi e la responsabilità legale sono temi centrali che richiedono una riflessione approfondita, basata su una visione multiprospettica, guidata dalla piena comprensione delle tecnologie e dal rispetto della dignità umana. Il concetto di "controllo umano significativo" deve quindi assumere un ruolo centrale nell'ambito delle ricerche sui sistemi d'arma intelligenti, al fine di evitare uccisioni indiscriminate, come purtroppo sta già accadendo oggi (Amoroso & Tamburrini, 2019a; 2020; Saidi, 2022).

## Conclusioni

L'utilizzo dei sistemi d'arma autonomi è un tema complesso, che coinvolge dimensioni etiche, giuridiche, filosofiche, politiche e militari. La stessa bibliografia esplorata in questo articolo riflette la diversità delle prospettive, evidenziando la necessità di un dialogo interdisciplinare per affrontare le sfide e garantire un utilizzo etico e responsabile di tali tecnologie. Il rapido sviluppo dell'IA in ambito militare ha profondamente trasformato la natura della guerra, introducendo nuove sfide e opportunità, ma anche molte preoccupazioni. Infatti, se da un lato l'implementazione dell'IA nei sistemi d'arma autonomi offre indiscussi vantaggi in termini di efficienza, rapidità e precisione, dall'altro è irrinunciabile affrontare le implicazioni etiche e stabilire regole chiare al fine di guidare correttamente il suo utilizzo in ambito militare. La guerra degli algoritmi è una realtà in continua evoluzione, e comprendere appieno il suo impatto è essenziale per plasmare il futuro delle dinamiche belliche globali prima che sia troppo tardi. Il dialogo internazionale e la cooperazione tra i paesi sono essenziali per garantire che l'IA contribuisca a un contesto di sicurezza globale anziché minacciarlo. La mancanza di regole chiare e il rischio di un utilizzo improprio delle tecnologie avanzate in ambito militare possono portare a conseguenze disastrose per l'intera umanità. La comunità internazionale è chiamata a sviluppare regole nuove, in grado di mitigare gli impatti negativi e garantire un uso responsabile dell'IA in guerra. Le nuove battaglie si svolgeranno quindi sul fronte dell'etica e della regolamentazione internazionale, nella speranza che in questo nuovo tipo di guerra non vi siano solo sconfitti.

## Bibliografia

Ackerman, E. (2014). U.S. Army Considers Replacing Thousands of Soldiers with Robots. *IEEE Spectrum*. <https://spectrum.ieee.org/army-considers-replacing-thousands-of-soldiers-with-robots>

Adams, T. K. (2001). Future Warfare and the Decline of Human Decisionmaking. *Parameters*, 31(4), 57-71.  
<https://press.armywarcollege.edu/cgi/viewcontent.cgi?article=2058&context=parameters>

Alexandre, L. (2018). *La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana*. Torino: EDT.

Allenby, B.R. (2015). *The Applied Ethics of Emerging Military and Security Technologies*. London: Routledge.

Allinson, J. (2015). The Necropolitics of Drones. *International Political Sociology*, 9.  
<https://doi.org/10.1111/ips.12086>

Amoore, L. (2009). Algorithmic war: Everyday geographies of the war on terror. *Antipode*, 41(1), 49-69.

Amoroso, D., & Tamburrini, G. (2019a). *What makes human control over weapons systems "meaningful"?*. ICRAC International Committee for Robot Arms Control. Report to the CCW GGE, Working Papers Series #4.

Amoroso, D., & Tamburrini, G. (2019b). I sistemi robotici ad autonomia crescente tra etica e diritto: quale ruolo per il controllo umano?. *BioLaw Journal – Rivista di BioDiritto*, 1, 33-51.

- Amoroso, D., & Tamburrini, G. (2020). Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues. *Current Robotics Reports*, 1, 187-194.
- Annati, M. (2019). Autonomous Weapon Systems. *Rivista Italiana di Difesa*, 4, 63-75.
- Anzarouth, M. (2021). Robots that Kill: The Case for Banning Lethal Autonomous Weapon Systems. *Harvard Political Review*, December 2. <https://harvardpolitics.com/robots-that-kill-the-case-for-banning-lethal-autonomous-weapon-systems/>
- Arkin, R. C. (2007). *Governing lethal behavior: embedding ethics in a hybrid deliberative/reactive robot architecture*. Technical Report GUT-GVU-07-11. Georgia Institute of Technology. <https://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf>
- Arkin, R. C. (2009). *Governing Lethal Behavior in Autonomous Robots*. New York, NY - Boca Raton, FL: Chapman and Hall – CRC Press.
- Armin, K. (2009). *Killer robots: Legality and ethicality of autonomous weapons*. Farnham, UK: Ashgate Publishing.
- Arquilla, J., & Ronfeldt, D. (1997). *In Athena's Camp: Preparing for Conflict in the Information Age*. Santa Monica, CA: Rand Corporation.
- Arquilla, J., & Ronfeldt, D. (2000). *Swarming and the Future of Conflict*. Santa Monica, CA: RAND Corporation.
- Asaro, P. (2008). How just could a robot war be?. In A. Briggle, K. Waelbers, & P. A. E. Brey (Eds.), *Current Issues in Computing And Philosophy (Frontiers in Artificial Intelligence and Applications)* (pp. 50-64). Amsterdam: IOS Press.
- Asaro, P. (2012). On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making. *International Review of the Red Cross*, 94(886), 687-709.
- Asaro, P. (2020). Autonomous Weapons and the Ethics of Artificial Intelligence. In S. M. Liao (Ed.), *Ethics of Artificial Intelligence*. Oxford: Oxford University Press.
- Asimov, I. (2021). *I, Robot*. Milano: Mondadori. (Edizione originale pubblicata nel 1950).
- Balistreri, M. (2015). Umanità e integrità del soldato potenziato: alcune riflessioni di bioetica militare. *Etica & Politica / Ethics & Politics*, XVII(1), 127-146.
- Balistreri, M. (2017). Robot Killer. La Rivoluzione Robotica nella Guerra e le Questioni Morali. *Etica & Politica / Ethics & Politics*, XIX(2), 405-430.
- Bandura, A. (2000). *Autoefficacia: teoria e applicazioni*. Trento: Erickson.
- Battaglia, P. (2018). *Intelligenza artificiale: dagli automi ai robot «intelligenti»*. Milano: RCS MediaGroup.
- Benjamin, M., & Ehrenreich, B. (2013). *Drone Warfare: Killing By Remote Control*. London: Verso Books.
- Bergen, P. L., & Rothenberg, D. (2014). *Drone Wars: Transforming Conflict, Law, and Policy*. New York, NY: Cambridge University Press.
- Bode, I., & Watts, T. (2023). *Loitering Munitions and Unpredictability. Autonomy in Weapon Systems and Challenges to Human Control*. University of Southern Denmark, Odense, DK: Center for War Studies – AutoNorms.

- Bode, I. (2024). Emergent Normativity: Communities of Practice, Technology, and Lethal Autonomous Weapon Systems. *Global Studies Quarterly*, 4, 1-11.
- Boulanin, V., Bruun, L., & Goussac, N. (2017). *Autonomous weapon systems and International Humanitarian Law: Identifying limits and the required type and degree of human-machine interaction*. SIPRI Stockholm International Peace Research Institute. [https://www.sipri.org/sites/default/files/2021-06/2106\\_aws\\_and\\_ihl\\_0.pdf](https://www.sipri.org/sites/default/files/2021-06/2106_aws_and_ihl_0.pdf)
- Boulanin, V., & Verbruggen, M. (2017). *Mapping the development of autonomy in weapon systems*. SIPRI Stockholm International Peace Research Institute. [https://www.sipri.org/sites/default/files/2017-11/siprireport\\_mapping\\_the\\_development\\_of\\_autonomy\\_in\\_weapon\\_systems\\_1117\\_1.pdf](https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_1.pdf)
- Boulanin, V., Davison, N., Goussac, N., & Carlsson, M. P. (2020). *Limits on autonomy in weapon systems. Identifying Practical Elements of Human Control*. SIPRI Stockholm International Peace Research Institute. [https://www.sipri.org/sites/default/files/2020-06/2006\\_limits\\_of\\_autonomy\\_0.pdf](https://www.sipri.org/sites/default/files/2020-06/2006_limits_of_autonomy_0.pdf)
- Bousquet, A. (2009). *The Scientific Way of Warfare: Order and Chaos on the Battlefields of Modernity*. New York: Columbia University Press.
- Burrige, B. (2003). *UAVs and the dawn of post-modern warfare: a perspective on recent operations*. *RUSI Journal*, 148(5), 18–23.
- Calhoun, L. (2016). *We Kill Because We Can: From Soldiering to Assassination in the Drone Age*. London: Zed Books.
- Čapek, K. (2015). *R.U.R. Rossum's Universal Robots*. Venezia: Marsilio Editore. (Edizione originale pubblicata nel 1921).
- Centro Alti Studi per la Difesa (CASD) – Istituto di Ricerca e Analisi della Difesa (IRAD). (2022a) (a cura di M. Artoni). *L'impiego dell'intelligenza artificiale – Una trasformazione inevitabile. Evoluzione e stato dell'arte*. [https://www.difesa.it/SMD\\_/CASD/IM/CeMiSS/DocumentiVis/Ricerche\\_da\\_publicare/Pubblicate\\_nel\\_2022/Artoni\\_AQ\\_SMD\\_12.pdf](https://www.difesa.it/SMD_/CASD/IM/CeMiSS/DocumentiVis/Ricerche_da_publicare/Pubblicate_nel_2022/Artoni_AQ_SMD_12.pdf)
- Centro Alti Studi per la Difesa (CASD) – Istituto di Ricerca e Analisi della Difesa (IRAD). (2022b) (a cura di C. Catalano). *Evoluzione e stato dell'arte dello "SWARMING" e la sua relazione con le Multi-Domain Operations (MDOs): progressi tecnologici, ambiti di applicazione, punti di forza, vulnerabilità, opportunità, minacce*. [https://www.difesa.it/SMD\\_/CASD/IM/CeMiSS/DocumentiVis/Ricerche\\_da\\_publicare/Pubblicate\\_nel\\_2023/ricerca\\_AQSMD11\\_Catalano.pdf](https://www.difesa.it/SMD_/CASD/IM/CeMiSS/DocumentiVis/Ricerche_da_publicare/Pubblicate_nel_2023/ricerca_AQSMD11_Catalano.pdf)
- Centro Alti Studi per la Difesa (CASD) – Istituto Superiore di Stato Maggiore Interforze (ISSMI). (2022c). *I principi fondamentali nella condotta delle operazioni militari*. [https://www.difesa.it/SMD\\_/CASD/IM/ISSMI/Documents/Precorso\\_Diritto\\_Umanitario.pdf](https://www.difesa.it/SMD_/CASD/IM/ISSMI/Documents/Precorso_Diritto_Umanitario.pdf)
- Ceola, P. (2020). *Macchine Guerriere Autonome*. *Filosofia*, LXV(4), 51-62. <https://ojs.unito.it/index.php/filosofia/article/view/5078/4702>
- Ceola, P., & Gaza, C. R. (2013). *Non-human Warfare. Robot e cyborg tra postmoderno e postumano*. Collana SISM, 4. Torino: Società Italiana di Storia Militare. <https://www.societaitalianastoriamilitare.org/COLLANA%20SISM/Non%20Human%20Warfare.pdf>

- Chamayou, G. (2015). *Teoria del drone. Principi filosofici del diritto di uccidere*. Roma: DeriveApprodi.
- Cheater, J. C. (2005). *Accelerating the kill chain via future unmanned aircraft*. Blue Horizons Paper, Maxwell Air Force Base, Montgomery AL, Center for Strategy and Technology, Air War College. <https://apps.dtic.mil/sti/pdfs/ADA497575.pdf>
- Chella, A. (2023). Coscienza artificiale: l'ingrediente mancante per un'IA etica?. *I Quaderni di Agendadigitale.eu, speciale Intelligenza Artificiale, 14*, 28-35.
- Chiriatti, M. (2021). *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*. Roma: LUISS University Press.
- Chu, J. (2017). *Army Robotics in the Military*. SEI Blog, June 12, Carnegie Mellon University. <https://insights.sei.cmu.edu/blog/army-robotics-in-the-military/>
- Clapper, J., Young, J., Cartwright, J., & Grimes, J. (2007). *Unmanned systems roadmap 2007–2032*. Arlington VA: Pentagon, Office of the Secretary of Defense.
- Clark, J. W. (1963). *Telechirics for Operations in Hostile Environments*. Battelle Technical Review, 3-8.
- Clark, J. W. (1964). *Remote Control in Hostile Environments*. *New Scientist*, 22(389), 300-304.
- Congressional Research Service. (2018). *U.S. Ground Forces Robotics and Autonomous Systems (RAS) and Artificial Intelligence (AI): Considerations for Congress*. CRS Report, R45392, 1-43. <https://crsreports.congress.gov/product/pdf/R/R45392>
- Contin, S. (2022). *Armi autonome, nuove sfide per l'umanità. Introduction about autonomous weapon*. Università di Verona, Facoltà di Filosofia. [https://www.researchgate.net/publication/364202002\\_Armi\\_autonome\\_nuove\\_sfide\\_per\\_l'umanita\\_-\\_introduction\\_about\\_autonomous\\_weapon](https://www.researchgate.net/publication/364202002_Armi_autonome_nuove_sfide_per_l'umanita_-_introduction_about_autonomous_weapon)
- De Landa, M. (1996). *La guerra nell'era delle macchine intelligenti*. Milano: Feltrinelli.
- Docherty, B. L. (2012). *Losing Humanity: The Case against Killer Robots*. Human Rights Watch, Cambridge MA. <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots>
- Dresp-Langley, B. (2023). The weaponization of artificial intelligence: What the public needs to be aware of. *Frontiers in Artificial Intelligence* (6)1154184. <https://www.frontiersin.org/articles/10.3389/frai.2023.1154184/full>
- Ekelhof, M. (2019). *Moving Beyond Semantics on Autonomous Weapons: Meaningful Human Control in Operation*. *Global Policy*, 10(3), 343-348.
- ElMasry, E. (2018). *Army of the Future: Artificial Intelligence and Its Impact on Operations*. JCSP-PCEMI 44, Canadian Forces College. <https://www.cfc.forces.gc.ca/259/290/405/192/elmasry.pdf>
- Etzioni, A., & Etzioni, O. (2017). Pros and Cons of Autonomous Weapons Systems. *Military Review*, 3, 72-81. <https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/pros-and-cons-of-autonomous-weapons-systems.pdf>
- Felt, C. (2020). *Autonomous Weaponry Are Killer Robots in Our Future?*. University of Washington, The Henry M. Jackson School of International Relations. <https://jsis.washington.edu/news/autonomous-weaponry-are-killer-robots-in-our-future/>

- Fitzsimonds, J. R., & Mahnken, T. G. (2007). Military officer attitudes towards UAV adoption: Exploring institutional impediments to innovation. *JFQ Joint Force Quarterly*, 46, 96-103.
- Fossa, F., Schiaffonati, V., & Tamburrini, G. (2021). *Automi e persone. Introduzione all'etica dell'intelligenza artificiale e della robotica*. Roma: Carocci Editore.
- Freedman, L. (2017). *The Future of War: A History*. New York: PublicAffairs.
- Galeotti, M. (2023). *The Weaponization of Everything. A Field Guide to the New Way of War*. New Haven CT: Yale University Press.
- Galetta, G. (2023). Il modello Osint in ambito militare: dinamiche di open/closed access nel trattamento dell'informazione a fini strategici. *Rivista di Digital Politics*, 3(2), 413-442.
- Gillis, J. (2017). Warfighter Trust in Autonomy. *DSIAC Journal*, 4, 23-29.
- Gregory, D. (2011). The everywhere war. *The Geographical Journal*, 177, 238-250.
- Grondin, D. (2013). *War Beyond the Battlefield*. London: Routledge.
- Gusterson, H. (2016). *Drone: Remote Control Warfare*. Cambridge, MA: The MIT Press.
- Hagström, M. (2016). *Characteristics of Autonomous Weapon Systems*. In Expert Meeting: Autonomous Weapon Systems. Implications of Increasing Autonomy in the Critical Functions of Weapons (23-5). Geneva: ICRC International Committee of the Red Cross.
- Hambling, D. (2021). Drones May Have Attacked Humans Fully Autonomously for the First Time. *New Scientist*, May 27. <https://www.newscientist.com/article/2278852-drones-may-have-attacked-humans-fully-autonomously-for-the-first-time/>
- Heisbourg, F. (1999). *Il futuro della guerra*. Milano: Garzanti.
- Holmqvist, C. (2013). Undoing War: War Ontologies and the Materiality of Drone Warfare. *Millennium*, 41, 535-552.
- Horowitz, M. C. (2019). When Speed Kills: Lethal Autonomous Weapons Systems, Deterrence and Stability. *Journal of Strategic Studies*, 42(6), 764-788.
- Horowitz, M. C., & Scharre, P. (2015). *Meaningful Human Control in Weapon Systems: A Primer*. Washington, D.C.: CNAS Center for a New American Security. <https://www.cnas.org/publications/reports/meaningful-human-control-in-weapon-systems-a-primer>
- HRW – Human Rights Watch (2016). *Killer Robots and the Concept of Meaningful Human Control*. Memorandum to Convention on Conventional Weapons (CCW) Delegates. <https://www.hrw.org/news/2016/04/11/killer-robots-and-concept-meaningful-human-control>.
- Iaria, A. (2018). Da autonomi a completamente autonomi: l'applicazione dell'Intelligenza Artificiale nei sistemi d'arma autonomi (LAWS). *Rassegna della Giustizia Militare: Rivista della Giustizia e della Procedura Penale Militare*, 6, 24-32.
- ICRC – International Committee of the Red Cross. (2014). *Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects*. Expert Meeting Report. Geneva: ICRC. <https://www.icrc.org/en/document/report-icrc-meeting-autonomous-weapon-systems-26-28-march-2014>
- IRIAD – Istituto di Ricerche Internazionali Archivio Disarmo. (2020). LAWS: Lethal Autonomous Weapon Systems: La questione delle armi letali autonome e le possibili azioni italiane ed europee per un accordo internazionale. *IRIAD Review – Studi sulla pace e sui conflitti*, 7-8, 1-161.

<https://www.archiviodisarmo.it/view/K0Y2nX8-UWjKHNm9OQ83o96Kv0-oDTrYQW5IYIxM1dE/iriad-review-luglio-agosto.pdf>

Jaffer, J. (2016). *The Drone Memos: Targeted Killing, Secrecy, and the Law*. New York NY: The New Press.

Jean, C. (2006). La guerra virtuale. In C. Rastelli & G. Cerino Badone (Eds.), *Storia della guerra futura*. Roma: Società Italiana di Storia Militare.

Johnson, M. A. (2019). *Drones and the Ethics of Targeted Killing*. Oxford Research Encyclopedia of Communication.

Kaag, J., & Kreps, S. (2014). *Drone Warfare*. Cambridge: Polity Press.

Kahn, P. W. (2002). The Paradox of Riskless Warfare. *Philosophy and Public Policy Quarterly*, 22(3), 2-8.

[https://openyls.law.yale.edu/bitstream/handle/20.500.13051/2656/Paradox\\_of\\_Riskless\\_Warfare%20The.pdf?sequence=2&isAllowed=y](https://openyls.law.yale.edu/bitstream/handle/20.500.13051/2656/Paradox_of_Riskless_Warfare%20The.pdf?sequence=2&isAllowed=y)

Kallenborn, Z. (2021). *Swarm Talk: Understanding Drone Typology*. Modern War Institute, West Point NY. <https://mwi.usma.edu/swarm-talk-understanding-drone-typology/>

Klare, M. T. (2019). *Autonomous weapons systems and the Laws of War*. Arms Control Association, Washington D.C. <https://www.armscontrol.org/act/2019-03/features/autonomous-weapons-systems-laws-war>

Krieg, A., & Rickli, J. M. (2019). *Surrogate Warfare: The Transformation of War in the Twenty-First Century*. Washington, DC: Georgetown University Press.

Krieger, D. (2018). *Geospatial Intelligence: Understanding the Basics*. London: Routledge.

Krishnan, A. (2009). *Killer robots. Legality and ethicality of autonomous weapons*. Farnham UK: Ashgate Publishing Limited.

Kurzweil, R. (2022). *The Singularity is Nearer. When We Merge with AI*. New York. NY: Viking Press.

Layton, P. (2018). *Algorithmic Warfare: Applying Artificial Intelligence to Warfighting*. Canberra, AU: Air Power Development Centre.

Layton, P. (2021). *Fighting Artificial Intelligence Battles. Operational Concepts for Future AI-Enabled Wars* (Joint Studies Paper Series No. 4). Canberra, AU: ADC Publications - Australian Defence College, Centre for Defence Research,

Leonard, M. (2023). *L'era della non-pace. Perché la connettività porta al conflitto*. Milano: Bocconi University Press.

Lewis, D. (2020). *An Enduring Impasse on Autonomous Weapons*. Just Security. <https://www.justsecurity.org/72610/an-enduring-impasse-on-autonomous-weapons/>

Lin, P., Bekey, G., & Abney, K. (2008). *Autonomous military robotics: Risk, ethics, and design*. Arlington VA: U.S. Department of Navy, Office of Naval Research.

[https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1001&context=phil\\_fac](https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1001&context=phil_fac)

Lucas, G. R., Jr. (2013). Engineering, ethics & industry: The moral challenges of lethal autonomy. In B. Strawser (Ed.), *Killing by remote control: The ethics of an unmanned military* (pp. 211-228). Oxford: Oxford University Press.



- Maas, M. M. (2019). How viable is international arms control for military artificial intelligence? Three lessons from nuclear weapons. *Contemporary Security Policy*, 40(3), 285-311.
- Marazzi, A. (2012). *Uomini, cyborg e robot umanoidi: Antropologia dell'uomo artificiale*. Roma: Carocci.
- Marchant, G. E., Allenby, B.R., Arkin, R., Barrett, E. T., Borenstein, J., Gaudet, L. M., Kittrie, O., Lin, P., Lucas, G. R., O'Meara, R., & Silberman, J. (2011). International Governance of Autonomous Military Robots. *Columbia Science and Technology Law Review*, XII, 272-315. <https://doi.org/10.7916/D8TB1HDW>
- Marsili, M. (2021) *Military Emerging Disruptive Technologies: Compliance with International Law and Ethical Standards. Intelligent and Autonomous: Emergent Digital Technologies and the Challenges of Disinformation, Security and Regulation*. Kaunas, LT: Vytautas Magnus University. <http://doi.org/10.5281/zenodo.4587612>
- Meaker, M. (2023). Ukraine's War Brings Autonomous Weapons to the Front Lines. *Wired*, February 24. <https://www.wired.com/story/ukraine-war-autonomous-weapons-frontlines/>
- Mégret, F. (2012). War and the Vanishing Battlefield. *Loyola University Chicago International Law Review*, 9(1). <https://ssrn.com/abstract=1986548>
- Mies, G. (2010). Military robots of the present and the future. *AARMS*, 9(1), 125-137.
- O'Hanlon, M. E. (2000). *Technological Change and the Future of Warfare*. Washington D.C.: The Brookings Institution.
- Osoba, O. A., & Wesler, W. IV. (2017). *An Intelligence in Our Image – The Risks of Bias and Errors in Artificial Intelligence*. Santa Monica, CA: RAND.
- Otto, S., & Mănescu, G. (2023). Will Artificial Intelligence (AI) replace a Human Commander in the Army?. *Scientific Bulletin*, XXVIII, 1(55), 79-87.
- Panwar, R. S. (2020). *Lethal Autonomous Weapon Systems: Slaves not Masters! Meaningful Human Control, saving Lives and Non-Feasibility of a Pre-Emptive Ban*. Future Wars, 6 ottobre. <https://futurewars.rspanwar.net/lethal-autonomous-weapon-systems-slaves-not-masters-meaningful-human-control-saving-lives-and-non-feasibility-of-a-pre-emptive-ban/>.
- Payne, K. (2021). *I, Warbot: The Dawn of Artificially Intelligent Conflict*. Oxford: Oxford University Press.
- Pireddu, M., & Tursi, M. (a cura di). (2006). *Post-umano: Relazione tra uomo e tecnologia nella società delle reti*. Milano: Guerini & Associati.
- Pryer, D. A. (2013). The rise of the machines: Why increasingly “perfect” weapons help perpetuate our wars and endanger our nation. *Military Review*, 93(2), 14-24.
- Quaranta, P. (2021). Lo stato della robotica militare. *Rivista Italiana di Difesa*, 7, 64-72.
- Quarteroni, A. (2021). *Algoritmi per un nuovo mondo*. Bari: Edizioni Dedalo.
- Reeves, S., & Johnson, W. (2014). Autonomous weapons: Are you sure these are killer robots? Can we talk about it?. *The Army Lawyer*, 1, 25-31.
- Roff, H. M. (2014). The strategic robot problem: Lethal autonomous weapons in war. *Journal of Military Ethics*, 13(3), 211-227.

- Rosapepe, F. P. (2020). *Intelligenza artificiale. Limiti e pericoli. I.A. da guerra*. Lecce: Youcanprint.
- Royakkers, L., & Topolski, A. (2014). Military Robotics & Relationality: Criteria for Ethical Decision-Making. In J. van den Hoven, N. Doorn, T. Swierstra, B.-J. Koops, & H. Romijn (Eds.), *Responsible Innovation 1. Innovative Solutions for Global Issues*. Dordrecht, NL: Springer. doi: 10.1007/978-94-017-8956-1\_20.
- Ruschi, F. (2016). Il volo del drone: Verso una guerra post-umana? Una riflessione di Filosofia del Diritto Internazionale. *Jura Gentium*, 1, 1-28.
- Sabry, F. (2021). *Armi autonome. In che modo l'intelligenza artificiale prenderà il sopravvento sulla corsa agli armamenti*. Abu Dhabi, UAE: 1BK One Billion Knowledgeable.
- Sacchi, M. (2020). *La guerra delle macchine. Hacker, droni e androidi: perché i conflitti ad alta tecnologia potrebbero essere ingannevoli e terribilmente fatali*. Milano: Algama Editore.
- Saidi, I. (2022). The Weaponization of Artificial Intelligence in the Military: The Importance of Meaningful Human Control. *MAS Journal of Applied Sciences*, 7(2), 357-363.
- Sariel, Y. (Brigadier General Y.S.) (2021). *The Human-Machine Team: How to Create Synergy Between Human & Artificial Intelligence That Will Revolutionize Our World*. Tel Aviv, IL: Independently Published.
- Scharre, P. D. (2018). *Army of None. Autonomous Weapons and the Future of War*. New York, NY: W. W. Norton & Company.
- Scharre, P. D., & Horowitz, M. (2015). *An introduction to autonomy in weapons systems* (Working Paper). Washington D.C.: CNAS Center for New American Security.  
<https://www.cnas.org/publications/reports/an-introduction-to-autonomy-in-weapon-systems>.
- Scharre, P. D., & Seaboyer, A. (2021). *Human control of Autonomous Weapons*. Toronto: DRDC Defence Research and Development Canada. Available at [https://cradpdf.drdc-rddc.gc.ca/PDFS/unc405/p815744\\_A1b.pdf](https://cradpdf.drdc-rddc.gc.ca/PDFS/unc405/p815744_A1b.pdf).
- Schiaffonati, V., & Tamburrini, G. (2021). Armi autonome e controllo umano significativo. In F. Fossa, V. Schiaffonati, & G. Tamburrini (Eds.), *Automi e persone. Introduzione all'etica dell'intelligenza artificiale e della robotica*. Roma: Carocci Editore, 133-152.
- Schwarz, E. (2015). Prescription drones: On the techno-biopolitical regimes of contemporary “ethical killing”. *Security Dialogue*, 47. doi: 10.1177/0967010615601388.
- Seixas-Nunes, A. (2022a). *The Legality and Accountability of Autonomous Weapon Systems: A Humanitarian Law Perspective*. Cambridge: Cambridge University Press.
- Seixas-Nunes, A. (2022b). Introducing Autonomous Systems of War: The Challenges of Artificial Intelligence. In A. Seixas-Nunes, *The Legality and Accountability of Autonomous Weapon Systems: A Humanitarian Law Perspective* (pp. 9-44). Cambridge: Cambridge University Press.
- Seixas-Nunes, A. (2022c). Autonomous Weapon Systems and ‘Autonomy’: Weapons or Killer Robots? In A. Seixas-Nunes, *The Legality and Accountability of Autonomous Weapon Systems: A Humanitarian Law Perspective* (pp. 103-109). Cambridge: Cambridge University Press.
- Sharkey, A. (2019). Autonomous weapons systems, killer robots and human dignity. *Ethics and Information Technology*, 21(2), 75-87.

- Sharkey, N. E. (2012). The evitability of autonomous robot warfare. *International Review of the Red Cross*, 94(886), 787-799.
- Sharkey, N. E. (2016). Staying in the Loop: Human Supervisory Control of Weapons. In N. Bhuta, S. Beck, R. Geiss, H. Y. Liu & C. Kress (Eds.), *Autonomous Weapons Systems: Law, Ethics, Policy* (pp. 23-38). Cambridge: Cambridge University Press.
- Shaw, I. G. R. (2013). Predator Empire: The Geopolitics of U.S. Drone Warfare. *Geopolitics*, 18(3), 536-559.
- Shaw, I. G. R. (2016). *Predator Empire: Drone Warfare and Full Spectrum Dominance*. Minneapolis, MN: University of Minnesota Press.
- Shaw, I. G. R. (2017). Robot Wars: US Empire and geopolitics in the robotic age. *Security Dialogue*, 48(5), 451-470.
- Singer, P. W. (2009a). *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. New York, NY: Penguin Books Ltd.
- Singer, P. W. (2009b). Wired for War? Robots and Military Doctrine. *JFQ Joint Force Quarterly*, 52(1), 104-110. Available at: [http://intelros.ru/pdf/jfq\\_52/21.pdf](http://intelros.ru/pdf/jfq_52/21.pdf).
- Singer, P. W. (2010). La Guerra delle macchine. *Le Scienze*, 505(9), 80-87.
- Sisinna, G. (2022). Come l'intelligenza artificiale sta trasformando la sicurezza nazionale – parte I. *AI4Business* (3 novembre). Available at: <https://www.ai4business.it/intelligenza-artificiale/come-lintelligenza-artificiale-sta-trasformando-la-sicurezza-nazionale-parte-i/>
- Smith, E. A., Jr. (2001). Network-Centric Warfare. *Naval War College Review*, 54(1), 5. Available at <https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=2391&context=nwc-review>.
- Smith, J. (2020). Autonomous Weapon Systems: A Comprehensive Survey. *Journal of Military Ethics*, 19(3), 245-263.
- Sparrow, R. (2007). Killer robots. *Journal of Applied Philosophy*, 24(1), 62-77.
- Stauffer, B. (2020). *Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control*. New York, NY: HRW Publications.
- Strawser, B. J. (2010). Moral predators: the duty to employ uninhabited aerial vehicles. *Journal of Military Ethics*, 9(4), 342-368.
- Strawser, B. J. (Ed.). (2013). *Killing by Remote Control: The Ethics of an Unmanned Military*. Oxford: Oxford University Press.
- Sullins, J. (2010). RoboWarfare: Can robots be more ethical than humans on the battlefield?. *Ethics and Information Technology*, 12(3), 263-275.
- Tamburrini, G. (2020). *Etica delle macchine. Dilemmi morali per robotica e intelligenza artificiale*. Roma: Carocci Editore.
- Tognocchi, M. (2018). *Drone warfare and the metamorphosis of battlefield: security, space and technology*. MAIR Thesis Seminar, Leiden University, Leiden, NL. Available at <https://studenttheses.universiteitleiden.nl/access/item%3A2660055/view>.
- Umbrello, S. (2021). Coupling levels of abstraction in understanding meaningful human control of autonomous weapons: a Two-tiered approach. *Ethics and Information Technology*. doi: 10.1007/s10676-021-09588-w.

U.S. Army Capabilities Integration Center – Future Warfare Division (2018). *Operationalizing Robotic and Autonomous Systems in Support of Multi-Domain Operations* (White Paper). Available at <https://info.publicintelligence.net/USArmy-RoboticAutonomousMultiDomainOps.pdf>.

U.S. Army Training and Doctrine Command (2017). *The U.S. Army Robotic and Autonomous Systems Strategy*. Fort Eustis, VA. Available at [https://mronline.org/wp-content/uploads/2018/02/RAS\\_Strategy.pdf](https://mronline.org/wp-content/uploads/2018/02/RAS_Strategy.pdf).

U.S. Department of Defense (2023). *Autonomy in Weapons Systems, directive no. 3000.09*. Pentagon, Arlington, VA. Available at <https://s3.documentcloud.org/documents/23588964/300009-autonomy-in-weapons-systems.pdf>.

U.S. Department of the Army – TRADOC, Army Capabilities Integration Center (ARCIC), Maneuver, Aviation, and Soldier Division (2017). *The U.S. Army Robotic and Autonomous Systems Strategy*. Fort Eustis, VA: TRADOC United States Army Training and Doctrine Command.

U.S. Department of the Navy - United States Marine Corps, Warfighting Laboratory, Wargaming Division (2017). *Manned Unmanned Teaming Wargame - Final Report*. Quantico, VA: Marine Corps Warfighting Laboratory.

U.S. Department of State (2023). *Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy*. Available at: <https://www.state.gov/politicaldeclaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/>

Valavanis, K. P., & Vachtsevanos, G. J. (Eds.). (2015). *Handbook of Unmanned Aerial Vehicles*. Dordrecht, NL: Springer Reference.

Veruggio, G., & Operto, F. (2008). *Roboethics: Social and ethical implications of robotics*. In B. Siciliano & O. Khatib (Eds.), *Springer handbook of robotics*. Berlin: Springer, 1499-1524.

Vicente, J. (2017). Glimpses of Tomorrow: The Coming Age of Autonomous Air Warfare. *Nação e Defesa*, 146, 66-84.

Vincent, J. (2021). Have Autonomous Robots Started War?. *The Verge*, June 3. <https://www.theverge.com/2021/6/3/22462840/killer-robot-autonomous-drone-attack-libya-un-report-context>

Wallach, W., & Allen, C. (2008). *Moral Machines: Teaching Robots Right from Wrong*. New York, NY: Oxford University Press.

Wang, L., & Zhang, Y. (2021). Predictive Analytics in Military Decision Making: A Review. *IEEE Access*, 9, 27662-27674.

Watts, T., & Bode, I. (2021). Worried about the autonomous weapons of the future? Look at what's already gone wrong. *The Bulletin of the Atomic Scientists*. <https://thebulletin.org/2021/04/worried-about-the-autonomous-weapons-of-the-future-look-at-whats-already-gone-wrong/>

Wilcox, L. (2017). Embodying Algorithmic War: Gender, Race, and the Posthuman in Drone Warfare. *Security Dialogue*, 48(1), 11-28.

Williams, A. J. (2011). Enabling persistent presence? Performing the embodied geopolitics of the unmanned aerial vehicle assemblage. *Political Geography*, 30, 381-390.

Williams, A. P., & Scharre, P. D. (Eds.). (2015). *Autonomous Systems: Issues for Defence Policymakers*. Norfolk, VA: NATO, Headquarters Supreme Allied Commander Transformation (printed by NATO Communications and Information Agency, The Hague, NL).

Work, R., & Brimley, S. (2014). *20YY Preparing for War in the Robotic Age*. Center for New American Security. Available at <https://www.cnas.org/publications/reports/20yy-preparing-for-war-in-the-robotic-age>.



# L'effetto degli smartphone sulle funzioni esecutive di bambini e adolescenti: cosa dicono gli studi

L'uso dello smartphone ha un profondo impatto sulla salute mentale e fisica dei bambini e potrebbe avere impatto su alcune funzioni cognitive dette funzioni esecutive. È quindi necessario promuovere un uso consapevole dei media digitali attraverso strategie educative adeguate e ulteriori ricerche

Di Giulia Conte, Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma, Dario Esposito, Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma, Mauro Ferrara, Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma, Sara Romano, Unità Operativa di Neuropsichiatria Infantile, Dipartimento di Neuroscienze Umane, Sapienza Università di Roma

Strumenti onnipresenti della vita quotidiana, gli smartphone hanno trasformato profondamente le abitudini di comunicazione, intrattenimento e accesso alle informazioni.

Le trasformazioni della nostra quotidianità sono del tutto evidenti: molto meno chiari sono i dati, a oggi disponibili, sull'effetto di questi strumenti sulla salute fisica e mentale dei loro utilizzatori, in particolar modo se bambini o adolescenti.

## Smartphone in età evolutiva: numeri e tendenze

Uno sguardo ai numeri del fenomeno smartphone in età evolutiva: una recente analisi su bambini e ragazzi di 8-18 anni negli USA (*The Common Sense Census*, 2022) ha evidenziato un aumento esponenziale nel numero di minori possessori di smartphone nel corso degli ultimi anni.

In particolare, nella fascia dei più piccoli, se nel 2015 solo un bambino di 8 anni ogni dieci possedeva uno smartphone, nel 2021 il rapporto è diventato di un bambino ogni tre. La stessa indagine ha evidenziato come i preadolescenti (età 8-12 anni) trascorrono 5 ore al giorno in media davanti ai vari media digitali (le femmine circa 5 ore in media, i maschi 6,11 ore), con un tempo ancora maggiore relativo ai teenager (12-18 anni). Per

quanto riguarda una delle tipologie più diffuse di utilizzo dei media, ovvero l'uso dei social media, circa il 57-70% di tutti i bambini e ragazzi li utilizza per almeno 2 ore al giorno, con un incremento progressivo tra i più giovani, nell'età compresa tra 8 e 12 anni.

## Raccomandazioni internazionali sull'uso dello smartphone dai più piccoli

Questo massiccio aumento dell'uso degli schermi digitali tra i bambini ha suscitato accesi dibattiti tra le agenzie istituzionali per la promozione della salute infantile. Ciò si è recentemente tradotto in una serie di raccomandazioni rilasciate da parte di Enti quali l'American Academy of Pediatrics che scoraggia l'esposizione a qualunque tipo di schermo digitale nei bambini prima dei 2 anni di età e sostiene l'utilità di un limite massimo di 1 ora al giorno per i bambini dai 2 ai 5 anni.

In Italia, tra le raccomandazioni emerse, un registro analogo è stato adottato dalla Società Italiana di Pediatria in collaborazione con Meta e Fondazione Carolina onlus, che hanno diramato nel 2022 l'iniziativa "Connessioni Delicate", specificando limiti orari e tipologie raccomandate di utilizzo dei media digitali per bambini in diverse fasce d'età.

Queste raccomandazioni perseguono dei chiari obiettivi di protezione dei minori dai rischi connessi all'uso eccessivo di schermi digitali in età evolutiva, ma nella comunità scientifica il dibattito circa la stima degli effetti dei media sui vari domini del benessere infantile – tra cui il sonno, le funzioni esecutive, il benessere fisico e lo stile alimentare, il benessere psicologico e la salute mentale – è a dir poco aperto.

## L'impatto dello smartphone sulle funzioni esecutive

Una delle tematiche più controverse riguarda l'impatto degli smartphone su *alcune* funzioni cognitive in via di sviluppo: l'attenzione, la memoria, le capacità di pianificazione e di decisione strategica. Nel loro insieme, queste competenze prendono il nome di funzioni esecutive (*FE*): raggruppano processi cognitivi diversi, per lo più inconsapevoli, che agiscono costantemente al servizio dell'organizzazione e della regolazione del comportamento in base all'obiettivo che ci poniamo nell'eseguire un'azione, oppure alla necessità di adattare il nostro operato a quanto richiesto dalle situazioni esterne.

Le FE sono fondamentali per affrontare le attività di tutti i giorni, sono utilizzate di continuo negli scenari della vita quotidiana: organizzare lo studio per una verifica a scuola, programmare le attività da fare nel pomeriggio, oppure il dover stabilire modi alternativi per raggiungere casa se la metro viene chiusa per un guasto tecnico.

In età evolutiva, le FE si organizzano e si perfezionano seguendo le incredibili trasformazioni nello sviluppo cerebrale in questa fase di vita. Come per gran parte delle

funzioni cerebrali, le FE vanno incontro ad una maturazione protratta nel corso dello sviluppo infantile, con un consolidamento maggiore in età scolare-preadolescenziiale, al passo con la progressiva maturazione della corteccia prefrontale del nostro cervello (Wu et al., 2011; Xu et al., 2013).

Le FE sono più o meno sviluppate da persona a persona e sono estremamente influenzate dall'ambiente di crescita, da quanto cioè il bambino venga più o meno adeguatamente stimolato nel consolidamento di queste abilità (Best & Miller, 2010).

## Screen time e sviluppo delle funzioni esecutive nei bambini

Nell'esplorare gli effetti degli smartphone sulle FE in età evolutiva, la letteratura scientifica è stata dominata da studi che hanno preso in esame la questione del tempo di utilizzo degli schermi digitali, il cosiddetto *screen time*. Tra questi, alcuni studi hanno correlato uno *screen time* maggiore a peggiori prestazioni nelle FE in età prescolare (Li et al., 2022; McHarg et al., 2020; Nathanson et al., 2014; Paulich et al., 2021), altri hanno rilevato che è soprattutto il perdurare durante l'infanzia di uno *screen time* elevato, in contrapposizione a periodi isolati di elevato utilizzo, a determinare effetti deleteri sulle FE (Best, 2012; McHarg et al., 2020).

Diversi Autori hanno ipotizzato diversi possibili meccanismi attraverso i quali l'eccessiva esposizione ai media digitali potrebbe influire sullo sviluppo delle FE: la sostituzione ad attività più appropriate per lo sviluppo e l'apprendimento, l'aumentato rischio di esposizione a contenuti inappropriati, la destrutturazione del sonno notturno e il ridotto accesso ad esperienze di relazione/*caregiving* che promuovono e sostengono le FE (Kirkorian et al., 2008; Ponti, 2023; Radesky & Christakis, 2016).

Oltre agli studi di tipo comportamentale, che hanno cioè misurato le varie FE in bambini e adolescenti, abbiamo a disposizione sempre più dati di tipo morfo-funzionale, che riguardano gli aspetti di maturazione e modificazione dei *network* cerebrali coinvolti nelle FE in base allo *screen time*. Nei bambini di 3-5 anni, è stato osservato che uno *screen time* elevato si associa a una minore organizzazione delle aree cerebrali del linguaggio, dell'attenzione e dell'apprendimento (Hutton et al., 2020, 2022).

Da dati emersi da un altro studio su una vasta popolazione di bambini di 10 anni, è emerso che un maggiore *screen time* si associa nel 37% dei casi a modifiche in aree importanti per le funzioni cognitive di ordine superiore (corteccia prefrontale) (Paulus et al., 2019). Tuttavia, nell'interpretare queste modifiche, gli Autori del lavoro scientifico hanno enfatizzato come esse non siano ascrivibili in maniera unilaterale a processi di ritardata maturazione, quindi di ordine "negativo", ma che, anzi, per alcuni bambini, le modifiche osservate sono indicative di un aumento specializzato dell'organizzazione strutturale della corteccia cerebrale.

In sintesi, il rapporto tra FE e *screen time* è ancora complesso da decifrare e quando si guarda alla sintesi dei dati esistenti in letteratura, ovvero alle metanalisi, le evidenze



sono tutt'altro che conclusive. Da una recente metanalisi che racchiude i risultati di svariati studi comportamentali sulle FE in bambini di 0-6 anni sulla base dello *screen time* (Bustamante et al., 2023), non emerge infatti alcun ruolo significativo del tempo di utilizzo degli schermi digitali sullo sviluppo delle FE e sulla loro efficienza.

Appare pertanto evidente la necessità di ricorrere ad un approccio più multi-componenziale per aumentare la nostra capacità di comprendere quali modelli o profili di utilizzo degli schermi digitali siano più strettamente associati a conseguenze positive, negative o neutre sulle FE dei bambini. Inoltre, poiché la maggior parte dei genitori fatica ad aderire alle raccomandazioni sul tempo da far trascorrere ai bambini davanti allo schermo, l'utilizzo di un approccio più articolato all'uso dei dispositivi digitali sembra particolarmente utile per identificare quali usi siano da incoraggiare per promuovere esiti positivi per i bambini.

In tal senso, appare centrale l'identificazione delle circostanze in cui l'uso dei media digitali è favorevole alla salute e allo sviluppo, ampliando l'analisi al di là del fattore *screen time* e considerando aspetti quali il contenuto dei media ed il contesto del loro utilizzo.

Svariati studi hanno evidenziato che la fruizione tramite media digitali di programmi educativi adatti all'età infantile è positivamente associata alle competenze di FE, indipendentemente dal tempo trascorso davanti agli schermi (Huber et al., 2018; Yang et al., 2017; Zimmerman & Christakis, 2007).

In uno studio su varie tipologie di video di cartoni animati per bambini, sono emerse peggiori FE in bambini esposti prevalentemente a video di intrattenimento (ricchi di stimolazioni sensoriali e di frequenti transizioni nelle trame o nelle scene, che catturano l'attenzione ma che riducono le capacità di tolleranza dell'attesa o di tempi più lenti di svolgimento della narrazione) rispetto a quelle di bambini esposti a video con contenuti più educativi (Fan et al., 2022).

Pertanto, l'assenza di caratteristiche informative e di tempi di narrazione più "lenti" e quindi più consoni al consolidamento delle informazioni nei programmi di intrattenimento può porre i bambini in una modalità passiva di elaborazione delle informazioni quando guardano questo tipo di video (Bryant & Anderson, 1983), determinando risultati peggiori nei compiti di funzionamento esecutivo. Analogamente, altri studi su bambini e uso passivo dei social media, hanno evidenziato come tale utilizzo preannunci la tendenza nel bambino a scegliere ricompense più piccole ma immediate, rispetto a ricompense maggiori ma ritardate (*delay aversion*) e maggiormente in linea con la fatica dell'apprendimento nella vita quotidiana (Schulz van Endert & Mohr, 2022; Wegmann et al., 2020).

Risultati analoghi rispetto alle conseguenze negative sulle FE dell'uso passivo e di intrattenimento degli schermi digitali, in particolare dei social media, provengono anche da studi condotti con adolescenti. Un gruppo di ricerca coreano ha evidenziato come lo *scrolling* passivo e ripetitivo sui social, determini una riduzione dell'attività dei

network attentivi (network frontoparietale) e faccia entrare il cervello in modalità “stand-by” aumentando l’attività del network “di base” del cervello (*default mode network* – DMN), ovvero dell’insieme di circuiti neurali (corteccia prefrontale mediale e cingolo posteriore), che si attivano quando non si è impegnati a svolgere compiti specifici (Hu, Cui, et al., 2022).

I ricercatori hanno evidenziato come ciò non accada quando gli stessi adolescenti sono immersi nella lettura di un libro di fantascienza, ad indicare quanto un utilizzo passivo dei contenuti digitali possa avere un impatto significativo sulle funzioni cognitive (Hu, Cui, et al., 2022), anche dopo sole 4 settimane di aumentato utilizzo dei social media in questa modalità (Hu, Yu, et al., 2022). Altri dati confermano tale ipotesi, mostrando che gli adolescenti con uso problematico dei social (la cosiddetta “dipendenza da social media”) presentano alle neuroimmagini una minore capacità di attivare le aree cerebrali deputate all’orientamento volontario dell’attenzione verso una certa attività (*dorsal attention network* – DAN) (Lee et al., 2021).

## I fattori di rischio dei social media

Altri Autori hanno evidenziato come la struttura stessa dei social media costituisca un fattore di rischio per l’utilizzo passivo dei contenuti digitali, dati alcuni meccanismi propri della comunicazione e del passaggio delle informazioni su queste piattaforme (Bhargava & Velasquez, 2021; Giraldo-Luque et al., 2020), tra cui:

- l’eccesso di notifiche e messaggi di tipo visivo e uditivo, che possono comportare una cronica saturazione dei network attentivi, con frequente interruzione e dispersione dell’attenzione rivolta ad altre attività;
- il meccanismo dei “mi piace” cioè la conferma dell’accettazione sociale ricevuta dagli altri utenti, che genera una potenziale dipendenza dall’uso dei social;
- i contenuti personalizzati cioè contenuti audiovisivi e testuali individuati dagli algoritmi dei social che creano attivazione del network del piacere tramite rinforzi intermittenti e variabili. Questo si associa alla mancata presenza di segnali naturali di stop come l’orario sullo schermo o la necessità di ricaricare pagine successive possono portare a fenomeni di dipendenza in soggetti maggiormente predisposti;
- la FoMo (Fear of Missing out) cioè il costante bisogno di essere online per timore di perdere eventi, informazioni o contenuti scaturita dall’affiliazione tra utenti e dal bisogno di appartenenza che possono portare ad ansia da controllo, attivazione emotiva negativa e riduzione delle capacità di pensiero orientato alle attività contestuali.

## L’impatto del contesto d’uso degli schermi sulla salute dei bambini

Anche il contesto in cui il bambino utilizza gli schermi appare come un fattore importante da considerare. Si stima che quasi il 60% dei genitori nel Regno Unito e a Singapore espongano i bambini all'uso di schermi durante i pasti (Goh & Jacob, 2012; Wright et al., 2007). Secondo vari studi disponibili (Fruh et al., 2021; Neshteruk et al., 2021; Pearson et al., 2018), tale contesto di utilizzo degli schermi contribuisce a un maggiore apporto calorico e a una durata prolungata del pasto, oltre che a un rischio maggiore di avere un BMI (*body mass index*) nel *range* degli obesi nei bambini piccoli (Teekavanich et al., 2022). Altre ricerche suggeriscono che il ricorso agli schermi come mezzo per regolare il comportamento dei bambini possa avere un impatto negativo sul loro sviluppo delle capacità di autoregolazione (Azevedo et al., 2022; Radesky, Kistin, et al., 2016; Radesky, Peacock-Chambers, et al., 2016). Ad esempio, se un bambino fa i capricci in uno spazio pubblico, i genitori possono dargli un dispositivo portatile per aiutarlo a distrarlo dalla sua esperienza di emozioni negative, ma l'uso di strategie di regolazione delle emozioni basate sullo schermo è stato collegato allo sviluppo di un uso più problematico dei media (ad esempio, perdita di interesse per attività che non coinvolgono gli schermi) nei bambini piccoli (Coyne et al., 2021).

## Verso un uso consapevole dei media digitali: il futuro di ricerca e educazione

In conclusione, la possibile correlazione tra uso di media digitali e funzioni esecutive si rivela un argomento complesso e in continua evoluzione. I dati raccolti fino a oggi dimostrano che esiste un rapporto, influenzato e orientato da una serie di variabili che vanno al di là del semplice tempo totale di utilizzo. Alcune ricerche suggeriscono una correlazione negativa tra l'uso eccessivo dei media digitali e le funzioni esecutive, altre sembrano suggerire che alcune attività digitali possano persino migliorarle. Il mondo scientifico su questo argomento è tutt'altro che unanime: sono sul campo diversi punti di vista e risultati controversi, il che sottolinea la necessità di un approfondimento e avanzamento della ricerca.

In un'ottica di riduzione del danno, selezionare attentamente le attività di utilizzo dello schermo in base alle caratteristiche proprie dei bambini (ad esempio, età, interessi, profilo comportamentale) e in base alle caratteristiche del contenuto, alla natura dell'uso e al contesto d'uso potrà aiutare i genitori a fare scelte che massimizzino i benefici e minimizzino i danni derivanti dall'uso degli schermi digitali.

Il futuro della ricerca in questo ambito riguarderà lo studio più approfondito dell'impatto specifico dei vari tipi di media e format digitali, considerando le modalità di utilizzo (passivo/attivo), l'età, il contesto e altri fattori influenti, non ultimo l'educazione al buon uso dei media digitali all'interno delle famiglie.

## Conclusioni

Dovremo continuare a interrogarci su come sfruttare al meglio i media digitali per il potenziamento delle funzioni esecutive e per favorire strategie educative che orientino al meglio l'inarrestabile – e irrinunciabile – rapporto di bambini e adolescenti con la tecnologia.

## Bibliografia

Azevedo, E. C., Riter, H. da S., Pieta, M. A. M., & Frizzo, G. B. (2022). Digital Media use on Interactions Between Mother and Child: Differences in Infants' Early Years. *Paidéia (Ribeirão Preto)*, 32, e3210. <https://doi.org/10.1590/1982-4327e3210>

Best, J. R. (2012). Exergaming immediately enhances children's executive function. *Developmental Psychology*, 48(5), 1501–1510. <https://doi.org/10.1037/a0026648>

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>

Bhargava, V. R., & Velasquez, M. (2021). Ethics of the Attention Economy: The Problem of Social Media Addiction. *Business Ethics Quarterly*, 31(3), 321–359. <https://doi.org/10.1017/beq.2020.32>

Bryant, J., & Anderson, D. R. (1983). *Children's Understanding of Television: Research on Attention and Comprehension*. Academic Press, Inc.

Bustamante, J. C., Fernández-Castilla, B., & Alcaraz-Iborra, M. (2023). Relation between executive functions and screen time exposure in under 6 year-olds: A meta-analysis. *Computers in Human Behavior*, 145, 107739. <https://doi.org/10.1016/j.chb.2023.107739>

Coyne, S. M., Shawcroft, J., Gale, M., Gentile, D. A., Etherington, J. T., Holmgren, H., & Stockdale, L. (2021). Tantrums, toddlers and technology: Temperament, media emotion regulation, and problematic media use in early childhood. *Computers in Human Behavior*, 120, 106762. <https://doi.org/10.1016/j.chb.2021.106762>

Fan, L., Lu, M., Qi, X., & Xin, J. (2022). Do Animations Impair Executive Function in Young Children? Effects of Animation Types on the Executive Function of Children Aged Four to Seven Years. *International Journal of Environmental Research and Public Health*, 19(15), 8962. <https://doi.org/10.3390/ijerph19158962>

Fruh, S., S, W., K, H., C, H., Gm, H., S, S., Rj, G., H, H., & J, B. (2021). A practical approach to obesity prevention: Healthy home habits. *Journal of the American Association of Nurse Practitioners*, 33(11). <https://doi.org/10.1097/JXX.0000000000000556>

Giraldo-Luque, S., Afanador, P., & Fernández-Rovira, C. (2020). The Struggle for Human Attention: Between the Abuse of Social Media and Digital Wellbeing. *Healthcare*, 8, 497. <https://doi.org/10.3390/healthcare8040497>

Goh, D. Y., & Jacob, A. (2012). Perception of picky eating among children in Singapore and its impact on caregivers: A questionnaire survey. *Asia Pacific Family Medicine*, 11(1), 5. <https://doi.org/10.1186/1447-056X-11-5>

Hu, B., Cui, Y.-L., Yu, Y., Li, Y.-T., Yan, L.-F., Sun, J.-T., Sun, Q., Zhang, J., Wang, W., & Cui, G.-B. (2022). Combining Dynamic Network Analysis and Cerebral Carryover Effect to Evaluate the Impacts of Reading Social Media Posts and Science Fiction in the Natural State on the Human Brain. *Frontiers in Neuroscience*, 16. <https://www.frontiersin.org/articles/10.3389/fnins.2022.827396>

Hu, B., Yu, Y., Yan, L., Qi, G., Wu, D., Li, Y., Shi, A., Liu, C., Shang, Y., Li, Z., Cui, G., & Wang, W. (2022). Intersubject correlation analysis reveals the plasticity of cerebral functional connectivity in the long-term use of social media. *Human Brain Mapping*, 43(7), 2262–2275. <https://doi.org/10.1002/hbm.25786>

Huber, B., Yeates, M., Meyer, D., Fleckhammer, L., & Kaufman, J. (2018). The effects of screen media content on young children's executive functioning. *Journal of Experimental Child Psychology*, 170, 72–85. <https://doi.org/10.1016/j.jecp.2018.01.006>

Hutton, J. S., Dudley, J., DeWitt, T., & Horowitz-Kraus, T. (2022). Associations between digital media use and brain surface structural measures in preschool-aged children. *Scientific Reports*, 12, 19095. <https://doi.org/10.1038/s41598-022-20922-0>

Hutton, J. S., Dudley, J., Horowitz-Kraus, T., DeWitt, T., & Holland, S. K. (2020). Associations Between Screen-Based Media Use and Brain White Matter Integrity in Preschool-Aged Children. *JAMA Pediatrics*, 174(1), e193869. <https://doi.org/10.1001/jamapediatrics.2019.3869>

Kirkorian, H. L., Wartella, E. A., & Anderson, D. R. (2008). Media and young children's learning. *The Future of Children*, 18(1), 39–61. <https://doi.org/10.1353/foc.0.0002>

Lee, D., Lee, J., Namkoong, K., & Jung, Y.-C. (2021). Altered functional connectivity of the dorsal attention network among problematic social network users. *Addictive Behaviors*, 116, 106823. <https://doi.org/10.1016/j.addbeh.2021.106823>

Li, S., Guo, J., Zheng, K., Shi, M., & Huang, T. (2022). Is Sedentary Behavior Associated With Executive Function in Children and Adolescents? A Systematic Review. *Frontiers in Public Health*, 10, 832845. <https://doi.org/10.3389/fpubh.2022.832845>

McHarg, G., Ribner, A. D., Devine, R. T., & Hughes, C. (2020). Screen Time and Executive Function in Toddlerhood: A Longitudinal Study. *Frontiers in Psychology, 11*. <https://www.frontiersin.org/articles/10.3389/fpsyg.2020.570392>

Nathanson, A. I., Aladé, F., Sharp, M. L., Rasmussen, E. E., & Christy, K. (2014). The relation between television exposure and executive function among preschoolers. *Developmental Psychology, 50*(5), 1497–1506. <https://doi.org/10.1037/a0035714>

Neshteruk, C., A, Z., L, S., E, E., We, K., Js, L., Ac, S., M, S., N, Z., & Sc, A. (2021). Weight-Related Behaviors of Children with Obesity during the COVID-19 Pandemic. *Childhood Obesity (Print), 17*(6). <https://doi.org/10.1089/chi.2021.0038>

Paulich, K. N., Ross, J. M., Lessem, J. M., & Hewitt, J. K. (2021). Screen time and early adolescent mental health, academic, and social outcomes in 9- and 10- year old children: Utilizing the Adolescent Brain Cognitive Development <sup>SM</sup> (ABCD) Study. *PLoS ONE, 16*(9), e0256591. <https://doi.org/10.1371/journal.pone.0256591>

Paulus, M. P., Squeglia, L. M., Bagot, K., Jacobus, J., Kuplicki, R., Breslin, F. J., Bodurka, J., Sheffield Morris, A., Thompson, W. K., Bartsch, H., & Tapert, S. F. (2019). Screen Media Activity and Brain Structure in Youth: Evidence for Diverse Structural Correlation Networks from the ABCD Study. *NeuroImage, 185*, 140–153. <https://doi.org/10.1016/j.neuroimage.2018.10.040>

Pearson, N., Sjh, B., P, G., Jp, J., & E, H. (2018). Clustering and correlates of screen-time and eating behaviours among young children. *BMC Public Health, 18*(1). <https://doi.org/10.1186/s12889-018-5698-9>

Ponti, M. (2023). Screen time and preschool children: Promoting health and development in a digital world. *Paediatrics & Child Health, 28*(3), 184–202. <https://doi.org/10.1093/pch/pxac125>

Radesky, J. S., & Christakis, D. A. (2016). Increased Screen Time: Implications for Early Childhood Development and Behavior. *Pediatric Clinics of North America, 63*(5), 827–839. <https://doi.org/10.1016/j.pcl.2016.06.006>

Radesky, J. S., Kistin, C., Eisenberg, S., Gross, J., Block, G., Zuckerman, B., & Silverstein, M. (2016). Parent Perspectives on Their Mobile Technology Use: The Excitement and Exhaustion of Parenting While Connected. *Journal of Developmental and Behavioral Pediatrics: JDBP, 37*(9), 694–701. <https://doi.org/10.1097/DBP.0000000000000357>

Radesky, J. S., Peacock-Chambers, E., Zuckerman, B., & Silverstein, M. (2016). Use of Mobile Technology to Calm Upset Children: Associations With Social-Emotional Development. *JAMA Pediatrics, 170*(4), 397–399. <https://doi.org/10.1001/jamapediatrics.2015.4260>

Schulz van Endert, T., & Mohr, P. N. C. (2022). Delay Discounting of Monetary and Social Media Rewards: Magnitude and Trait Effects. *Frontiers in Psychology, 13*. <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.822505>

Teekavanich, S., Rukprayoon, H., Sutthritpongsa, S., & Rojmahamongkol, P. (2022). Electronic media use and food intake in Thai toddlers. *Appetite, 176*, 106121. <https://doi.org/10.1016/j.appet.2022.106121>

*The Common Sense Census: Media Use by Tweens and Teens, 2021 | Common Sense Media*. (n.d.). Retrieved November 21, 2023, from <https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2021>

Wegmann, E., Müller, S. M., Turel, O., & Brand, M. (2020). Interactions of impulsivity, general executive functions, and specific inhibitory control explain symptoms of social-networks-use disorder: An experimental study. *Scientific Reports, 10*(1), Article 1. <https://doi.org/10.1038/s41598-020-60819-4>

Wright, C. M., Parkinson, K. N., Shipton, D., & Drewett, R. F. (2007). How do toddler eating problems relate to their eating behavior, food preferences, and growth? *Pediatrics, 120*(4), e1069-1075. <https://doi.org/10.1542/peds.2006-2961>

Wu, K. K., Chan, S. K., Leung, P. W. L., Liu, W.-S., Leung, F. L. T., & Ng, R. (2011). Components and developmental differences of executive functioning for school-aged children. *Developmental Neuropsychology, 36*(3), 319–337. <https://doi.org/10.1080/87565641.2010.549979>

Xu, F., Han, Y., Sabbagh, M. A., Wang, T., Ren, X., & Li, C. (2013). Developmental differences in the structure of executive function in middle childhood and adolescence. *PloS One, 8*(10), e77770. <https://doi.org/10.1371/journal.pone.0077770>

Yang, X., Chen, Z., Wang, Z., & Zhu, L. (2017). The Relations between Television Exposure and Executive Function in Chinese Preschoolers: The Moderated Role of Parental Mediation Behaviors. *Frontiers in Psychology, 8*, 1833. <https://doi.org/10.3389/fpsyg.2017.01833>

Zimmerman, F. J., & Christakis, D. A. (2007). Associations between content types of early media exposure and subsequent attentional problems. *Pediatrics, 120*(5), 986–992. <https://doi.org/10.1542/peds.2006-3322>

# Ri-organizzare la Realtà Estesa: il ruolo flessibile della cultura e delle simulazioni filosofiche e tecnologiche

In che modo epoché filosofica, interrogativi controintuitivi, condizionali controfattuali e rappresentazioni, tecnologia inclusa, influenzano il nostro rapporto con la realtà, permettendone una riorganizzazione flessibile? La cultura funge da semplificatore, mentre filosofia, arte e tecnologie come AR/VR e robot interrompono le ripetizioni compulsive

Di Carmine T. Recchiuto, Università degli Studi di Genova, DIBRIS, RICE Lab, Lorenza Saettone, Filosofa specializzata in Epistemologia e Cognitivism, PhD Student in Robotics and Intelligent Machines for Healthcare and Wellness of Persons e Antonio Sgorbissa, Università degli Studi di Genova, DIBRIS, RICE Lab

La filosofia, nel corso della storia, molto spesso ha posto in questione le cosiddette “posizioni predefinite”<sup>17</sup>, quelle cioè che fanno da sfondo a ogni condizione del vivere, del pensare e del comunicare. Il dubbio si è scagliato, in primo luogo, sulla realtà e, in seconda istanza, sulla possibilità di conoscerla e di comunicarla.

Il motivo del mettere tra parentesi l’essere e la sua dicibilità dipende dal fatto di avere esperienza solo del Fenomeno, cioè della nostra percezione soggettiva, mentre la Cosa in Sé, il contenuto del pensiero a prescindere dal pensiero stesso, resta inevitabilmente un’incognita *x*, intuibile, ma non conoscibile. Quest’ultimo punto di pensabilità-impensabile fu inteso dai post-kantiani come una contraddizione. Per conservare la consistenza del sistema, la *x* venne perciò eliminata, mantenendo solo il soggetto, che, oltre all’idea, porrebbe anche il contenuto della stessa. Ecco l’idealismo. Searle<sup>18</sup> addita questo atteggiamento, considerandolo il frutto di una volontà di potenza in risposta al senso di inferiorità delle scienze dello spirito. Se tutto è idea, la scienza deve appoggiarsi a chi è maestro di idee: gli umanisti.

Anche il dubbio metodologico (epoché) intorno ai presupposti, a mio avviso non è solo figlio di un bisogno di potere e quindi di una sorta di morale del risentimento. I presupposti sono stati negati o posti in dubbio solo in istanza filosofica, al contrario, nell’agire concreto, sono sempre stati dati per scontati. Dubitare dell’esistenza del pane, del corpo, e di ogni altra condizione di sopravvivenza ha senso fino a che non iniziano la fame e altre urgenze biologiche; lì, il dubbio lascia inevitabilmente spazio alla Vita.

In questa sede ci si domanda quale sia il significato del dubbio e della filosofia, anche in un’ottica evolutiva, dato che la loro attività sembra prima facie in contrasto con il concetto di utile pragmatico. Lo stesso si dica dell’arte, la quale risulterebbe inutile se l’unico criterio di giudizio fosse la sopravvivenza immediata, la risposta deterministica in un ambiente esteso

<sup>17</sup> J. Searle, *Mente, Linguaggio e Società*, Scienze e idee, Raffaello Cortina Editore, Milano, 2000.

<sup>18</sup> *Ibidem*.



privo di evoluzione. Sia arte sia filosofia hanno funzioni differenti in risposta a motivi diversi<sup>19</sup>. Entrambe creano un attrito, un'attesa all'agire altrimenti riflesso ed è in questa attesa che si situa una possibilità di riadattamento e perfezionamento.

Il fatto di essere organismi significa che siamo organizzati secondo attività di vario genere a cui partecipiamo di continuo. In quest'ottica, l'attrito è un momento di ridefinizione dell'attività spontanea primigenia, nel quale quest'ultima si fa proprio oggetto di rappresentazione dell'arte e della filosofia. La coreografia rappresenta la danza spontanea, l'epistemologia rappresenta il pensare spontaneo intorno alla realtà. Da tali rappresentazioni riemergeranno, rispettivamente, movimenti e pensiero più calibrati<sup>20</sup>, senza ridondanze, che troveranno spazio di nuovo in quell'agire spontaneo, da cui arte e filosofia avevano preso le mosse e, contemporaneamente, le distanze. Non sono parte della Vita e utili nell'immediato; la loro utilità rappresenta un effetto secondario nel continuum noi-mondo-altri in cui inevitabilmente siamo posti.

Lo stesso si dica di ogni forma di rappresentazione del mondo. Le simulazioni non vogliono essere la sostituzione della Cosa in Sé, bensì metodi di comprensione e di ri-organizzazione dell'originale e di noi stessi rispetto all'originale. Lo stesso metodo con cui rappresentiamo, nell'arte, nella filosofia e nelle tecnologie, le azioni e i fatti del mondo serve agli esseri umani per allenare la costruzione delle istituzioni, applicando la formula searliana “Questo X vale come Y in C”<sup>21</sup>, cioè, tradotto in linguaggio naturale: “Questo pezzo di metallo vale come cinque centesimi in Europa”. Rappresentare porta altresì a esercitare la capacità di scelte su mondi possibili. Questi ultimi sono rappresentazioni alternative del mondo attuale, che si rendono necessari per valutare la verità di proposizioni cosiddette modali. La logica classica non è sufficiente a chiarire le proposizioni non vero-funzionali, la cui verità non è immediatamente comprensibile dalle sole tavole di verità. Nelle proposizioni in cui compaiono gli avverbi di tempo e gli atteggiamenti modali come “è necessario, è possibile, è obbligatorio, è permesso, è creduto, è saputo”, la verità viene valutata su mondi possibili, *i quali rappresentano una struttura completa di tutti i fatti che potrebbero essere stati sia uguali sia diversi dal mondo attuale*. Kripke<sup>22</sup> introdusse i mondi possibili per spiegare le verità necessarie e contingenti, proponendo che le affermazioni modalmente necessarie siano vere in tutti i mondi possibili, mentre le affermazioni contingenti possano variare da mondo a mondo.

Insomma, quando un agente può credere a una falsità, quando compaiono altri soggetti nelle descrizioni dei fatti, i mondi si moltiplicano tra rappresentazioni alternative. Vedremo più avanti come questa esplosione di mondi sia una euristica che semplifica la scelta e non una ridondanza che complica la realtà-una.

Insomma, come il dubbio a proposito della realtà non intende sostituire la partenza di cui si dubita, così la mappa non è il territorio e i mondi possibili non sono reali nel senso del realismo esterno, ma hanno conseguenze reali, causali, sulle azioni che poi vengono intraprese nel mondo attuale.

Il realismo esterno afferma l'esistenza della realtà indipendente da noi, cioè di sedie, banchi, luna, quark, i quali continuano ad essere anche quando non li percepiamo, ed è questo presupposto esterno da dare per scontato nell'agire. Tuttavia possiamo dare per scontato proprio perché non siamo un cervello in una vasca, proprio perché non sussiste il dubbio sul terreno dell'azione immediata. Se fossimo un algoritmo non potremmo paradossalmente dare per scontato alcunché: se voglio fare in modo che uno sprite di un videogame apra una porta, prima

<sup>19</sup> A. Noë. *Strange Tools: Art and Human Nature*. First edition. New York, Hill and Wang, a division of Farrar, Straus and Giroux, 2015.

<sup>20</sup> *Ibidem*

<sup>21</sup> J. Searle, *Mente, Linguaggio e Società*, Scienze e idee, Raffaello Cortina Editore, Milano, 2000.

<sup>22</sup> Kripke, *S* 1980 Naming and Necessity Cambridge Harvard University Press

devo definire l'esistenza della porta, del soggetto e della fisica per aprirla. Noi, invece, possiamo dimenticarci che esistono le cose, la gravità e porle solo in certe condizioni in dubbio perché sono presupposte e reali. Da ciò si comprende quanto la teoria della simulazione artificiale nel campo della fisica sia insostenibile<sup>23</sup>.

Non sono il realismo e i fatti bruti che meritano un'indagine, di questo se ne occupa la scienza. Quello di cui ci occupiamo è il motivo per il quale vengono messe in dubbio o simulate. Allo stesso modo ciò che ci interessa indagare è un altro tipo di dare per scontato, cioè quello che entra in gioco quando affermiamo che Napoleone era un uomo basso che non riguarda l'esistenza di un soggetto umano maschile di nome Napoleone e che questo avesse la proprietà di essere dotato di bassezza, ma il fatto di non dover chiarire di quale Napoleone si tratti, giacché tutti sappiamo di attingere dallo stesso bacino di presupposizioni culturali, sociali, educative, esperienziali, specifiche, le quali fanno in modo, quando dico Napoleone + basso, di dirigerci tutti al capitolo della storia che segue la Rivoluzione Francese.

Come viene costruito questo bacino di presupposti? Quale status ontologico ha? Che cos'è questo secondo tipo di dare per scontato? Cos'è la soggettività che dubita e che è capace di allacciarsi all'oggetto secondo diverse modalità e forze? Ha una natura irrealistica o è reale? È reale nel senso dei muoni o della liquidità? Ha ancora senso utilizzare le categorie proprie degli approcci dualistici e monistici? Giacché anche chi invoca l'onestà intellettuale materialistica o idealistica (entrambi credono di star rasoando con il gillette di Occam; eppure, il barbiere in questione finisce quasi sempre per rivelarsi quello di Russell<sup>24</sup>), usa i concetti del dualismo: materia o forma, idea o cosa, cervello o coscienza.

Noi qui proponiamo una forma di realismo ingenuo e di mente estesa<sup>25</sup>, comprensibile solo negli usi situati, collettivi e corporei, e materializzata e realizzata nelle e con le affordances<sup>26</sup>.

## Coscienza realizzata

Nella prospettiva Cartesiana, il dubbio utilizzato a mo' di metodo diventava presto ipertrofico, cioè viene applicato a qualunque aspetto soggettivo e oggettivo, potendo concepire l'esistenza di un genio maligno atto a ingannare l'essere umano sistematicamente e su tutto quanto.

Eppure, come mostrava lo stesso René Descartes, la totale messa tra parentesi faceva emergere l'Io come presupposto innegabile: Ego Dubito, Ego Cogito ergo Sum, poiché per dubitare deve

---

<sup>23</sup> Sisini, Francesco & Sisini, Valentina & Sisini, Laura. (2023). Limiti teorici della Virtual Reality. 10.13140/RG.2.2.27449.65126.

<sup>24</sup> *Paradosso del barbiere*. (2023, April 22). Wikipedia. [https://it.wikipedia.org/wiki/Paradosso\\_del\\_barbiere](https://it.wikipedia.org/wiki/Paradosso_del_barbiere)

<sup>25</sup> Il "realismo ingenuo" è una posizione filosofica che sostiene che il mondo esterno esista indipendentemente dalla nostra percezione o esperienza di esso. Questa visione postula che gli oggetti e gli eventi esistano in modo oggettivo e indipendente dalle nostre menti e che possiamo avere una conoscenza accurata del mondo attraverso la percezione sensoriale e la ragione (Cfr. Austin, John Langshaw. *Sense and sensibilia* (1959), Oxford: Oxford UP, 1964 ; Putnam, Hilary. *Mente, corpo, mondo*. Il Mulino, 2003 ). Il concetto di "mente estesa" è una prospettiva che sfida il dualismo mente-corpo tradizionale, suggerendo che la mente non sia confinata all'interno del cranio di un individuo, ma possa estendersi oltre il corpo per includere oggetti e ambienti esterni. Questa idea, sviluppata da filosofi come Andy Clark e David Chalmers (Andy Clark e David J. Chalmers, *The Extended Mind*, in *Analysis*, vol. 58, n. 1, gennaio 1998, pp. 7-9, DOI:10.2307/3328150) suggerisce che le tecnologie e gli strumenti esterni, come computer, smartphone o quaderni, possono fungere da estensioni della mente umana, contribuendo al processo cognitivo e alla formazione delle nostre esperienze e conoscenze.

<sup>26</sup> L'affordance è un concetto introdotto dallo psicologo James J. Gibson per descrivere le potenziali azioni o interazioni che un oggetto o un ambiente offre a un individuo. In altre parole, si riferisce alle caratteristiche di un oggetto o di un ambiente che suggeriscono o consentono determinati utilizzi o comportamenti. Le affordance dipendono dalla percezione dell'individuo e dalle sue capacità fisiche e cognitive (Cfr. James Jerome Gibson. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979).

esistere un Io che metta in atto quell'azione stessa del pensiero. Diversamente dall'idealismo gnoseologico alla Cartesio e alla Berkeley, il materialismo dubita dell'Io cosciente, della sostanza pensante. Dall'esperienza delle percezioni viene problematizzato l'elemento soggettivo e individuale, mantenendo solo il fatto osservabile e misurabile, insomma, ciò che è esterno e pubblico.

Per uscire da questa antinomia, la via potrebbe essere quella di lasciare perdere tale "coperta teoretica" sempre troppo corta, con cui si copre o la materia o l'Io, abbracciando altri paradigmi non dualistici. La messa tra parentesi e il sospetto sull'identità, sulla mente e sui legami intenzionali di questa con le cose non vanno intesi come teorie sostitutive ai presupposti. Per non incorrere in inganni, sembra utile impiegare categorie differenti per definire il modo in cui esiste la cosiddetta coscienza in relazione ai suoi contenuti. Giacché utilizzare le categorie della tradizione filosofica rischia di generare incomprensioni, come quella di chiamare "intelligenza" una rete neurale artificiale che computa semplicemente algoritmi statistici.

Torniamo ai realisti. Secondo loro il mondo è reale perché possiamo correggere gli errori, perché più individui nello spazio e nel tempo e più sensi concordano sui medesimi fatti (a parte alcune questioni residuali, su cui però non c'è mai stata intraducibilità) e perché la nostra stessa aspettativa è confermata (o rifiutata) nei piani ad agire in modo non caotico e stocastico. È più semplice accettare che esista un mondo esterno piuttosto che dubitarne. Al contrario, l'idealismo finisce sempre per chiamare in causa un deus ex machina, ad hoc. Tuttavia, il realismo di cui la semplicità si fa carico è quello che comprende fatti bruti insieme alla mente estesa. È solo in questo continuum che realizziamo la coscienza, nel senso sia di "attuazione" sia di "consapevolezza".

Come sostengono gli enattivisti<sup>27</sup>, non siamo confinati dentro di noi, ma distribuiti e dinamicamente estesi in diversi coinvolgimenti del mondo, che comprendono tecnologie, abilità e istituzioni (riassunti, concetti ausiliari delle singole attività individuali ma attuate con la credenza in un Noi)<sup>28</sup>. Il realismo generale tende a dimenticarsi delle menti e dei fatti dipendenti dagli osservatori; questo lascia troppo spazio a incomprensioni e all'idea che i non fatti bruti siano illusioni.

*Realismo generico*<sup>29</sup>: *a*, *b*, *c* e così via esistono, e il fatto che esistano e abbiano proprietà come *F-ness*, *G-ness* e *H-ness* è (a parte le banali dipendenze empiriche del tipo talvolta incontrato in vita quotidiana) indipendente dalle credenze, dalle pratiche linguistiche, dagli schemi concettuali di ciascuno e così via.

Ciò che si vuole sottolineare come rilevante all'interno di questa definizione di realismo è il fatto che oggetti e proprietà esistono, sono reali, ma indipendentemente dalle menti e quindi dalle convenzioni sociali che ne definiscono pratiche e usi concreti. Insomma, si tratta dell'oggetto in sé, con le sue proprietà primarie, senza che intervenga un soggetto, con i suoi dubbi e le condizioni di felicità o di soddisfazione<sup>30</sup>. Il mondo del realista esterno pare costituito

<sup>27</sup> L'enattivismo è una prospettiva filosofica e cognitiva che enfatizza il ruolo attivo dell'organismo nell'interazione con il mondo esterno. Questa prospettiva si oppone al dualismo tradizionale mente-corpo e sostiene che la cognizione sia fondamentalmente radicata nell'azione e nell'interazione diretta con l'ambiente circostante. (Cfr. Noe A., *Perché non siamo il nostro cervello*, 2010, Raffaello Cortina Editore, Milano)

<sup>28</sup> Popper, K.R., *The poverty of historicism*, London 1957 (tr. it.: *Miseria dello storicismo*, Milano 1975<sup>2</sup>).

<sup>29</sup> *Realism*, in *Stanford Encyclopedia of Philosophy*, 2019, December 13. <https://plato.stanford.edu/entries/realism/>

<sup>30</sup> Nella pragmatica non si parla di condizioni di verità (cioè le condizioni ambientali che rendono vera o falsa una proposizione), ma si parla di condizioni di felicità, condizioni che portano a buon fine un'azione. Una preghiera non è vera o falsa. Riguarda un altro metodo di verifica, che prende le mosse da ruoli sociali, obiettivi, intenzioni del parlante, effetti perlocutori dei soggetti coinvolti nell'azione linguistica, adattamenti mondo-parola e parola-mondo che incorporano forze diverse dall'asserzione e che riguardano più la performance che la descrizione di come è o come non è il mondo (Cfr. Austin J. I. (1962). *How to do things with words*. Oxford University Press).

da semplici fatti, scomponibili in atomi osservabili da una non-mente e collegati a un linguaggio formale non praticato, in cui le cose sono o vere o false. È, tuttavia, solo una parte della realtà; è anche questa una scelta, una semplificazione. Anch'essa insomma è una delle rappresentazioni.

Dal momento in cui la soggettività si palesa nel dominio del realista, oltre a epistemologia e ontologia, anche la logica si complica: non esistono più proposizioni il cui senso si mostra facilmente nelle *truth tables*. Fanno capolino le probabilità, le valutazioni su diversi parametri, in cui la verità diventa qualcosa che somiglia più all'utile relativo. Intervengono condizioni di felicità, per le quali un'azione non è vera o falsa, ma attraverso cui essa va a buon termine o meno, nelle quali importanti sono i ruoli sociali e le convenzioni co-create e co-credute. Compaiono, in breve, i mondi possibili su cui valutare conoscenze, credenze, possibilità, obblighi e cambi temporali<sup>31</sup> e quindi relazioni di accessibilità tra quei mondi, con cui si possono tradurre i concetti di empatia (entrare nei mondi degli altri agenti) e di introspezione (conoscere e valutare quello che si sa e che si crede e assumere un comportamento razionale sulla base di quello che ancora non si conosce ma si crede tale)<sup>32</sup>.

Insomma, più un agente è coinvolto con il suo ego e con quello degli altri, più le scelte logiche si fanno ricche di simboli, operatori modali, tempo, dinamicità e variabili intervenienti che possono modificare le attese in corso d'opera. Si complica l'insieme di predicati, includendo, obblighi, possibilità, avverbi temporali e operatori epistemici. Parallelamente, per gestire questa ricchezza di modalità, i fatti del mondo, cioè gli *a, b, c* del realista, si riducono, per permettere un ragionamento più rapido ed efficace, che richieda una minore esplosione combinatoria sui mondi possibili. L'agente pre-esclude tutti quegli elementi di cui l'attenzione può fare a meno, tuttavia, perché ciò sia possibile deve sussistere un'attenzione preventiva e precosciente, che setacci ciò che è utile e ciò che non lo è. Come riesce ad attenzionare una riduzione del reale non potendo gestire la sua complessità? È un azzardo? Eppure quotidianamente constatiamo come ciò funzioni.

Se lo simulassimo al computer, avremmo bisogno di un programmatore che costruisca l'ontologia pensata appositamente per l'obiettivo; traslando questo nell'essere umano dovremmo analogamente ipotizzare un omuncolo, con un regresso all'infinito. Oppure, per risolvere la preselezione, escludendo il paradosso del terzo uomo, si fa appello a un intelletto sempre attivo, di aristotelica memoria, che deve in qualche modo aver già elaborato la complessità per noi. Per "intelletto attivo" si intende non un intelletto divino o un'anima con idee innate pascolate nell'Iperuranio, ma un bacino di presupposizioni sociali, risultato della testimonianza collettiva, sempre aperto a modifiche in itinere, proprio perché agito individualmente e flessibile. Nasciamo con la capacità di vocalizzare in un gruppo che partecipa già all'istituzione linguistica (tra le altre istituzioni); pian piano impariamo anche noi a emettere vocalizzi specifici e a gestirli nelle relazioni, secondo le intenzionalità più proprie. Il linguaggio ci precede, è da sempre attivo, e decidiamo di abitarlo a nostra volta perché ovviamente ci è utile<sup>33</sup>.

<sup>31</sup> Frixione M., Iaquinto S., Vignolo M., (2016) Introduzione alle logiche modali, Laterza, Roma Bari.

<sup>32</sup> L. Saettone, R. Fedriga, E. Micheli, (2023), Linguistic and Cultural Competencies in Dynamic Possible Worlds, RobOntics 2023 Ontologies for Autonomous Robotics 2023 Proceedings of the Workshop on Ontologies for Autonomous Robotics co-located with The 32nd IEEE International Conference on Robot and Human Interactive Communication (IEEE RO-MAN 2023) Busan, S. Korea, August 28, 2023.

<sup>33</sup> Tomasello, M. (2009, June 30). *Constructing a Language*. Harvard University Press.

*“Ciò che stiamo trascurando è il fatto che la coscienza non è solo una importante caratteristica della realtà. In un certo senso essa è la più importante perché tutte le cose hanno valore, rilevanza, merito, pregio soltanto in relazione alla coscienza”<sup>34</sup>.*

Searle sottolinea qui due aspetti fondamentali del realismo esteso da noi caldeggiato. Innanzitutto, una coscienza parte della realtà, superando in questo modo le categorie filosofiche della tradizione occidentale, con cui inevitabilmente si cade nei dualismi. Inoltre, la coscienza diventa ciò che, tramite l'intenzionalità, spiega la preselezione degli elementi utili per ulteriori scelte. Ogni essere umano usa la coscienza per relazionarsi al dominio di base (lo Sfondo di Searle), altrimenti ridondante di significati, selezionando, con le euristiche concettuali più utili in ogni dato momento, porzioni del dominio. È come se l'esperienza personale di ognuno, ma vissuta in un contesto sociale e pubblico, creasse una reazione, più o meno riflessa, di connessioni, pesi sinaptici e funzioni soglia con cui richiamare i pattern di attivazione in risposta a ogni input. Solo così le cose acquisiscono valore, e solo così, sulla base di quel valore (il peso), si possono selezionare dalla memoria semantica e procedurale i piani migliori. È questo che chiamiamo azione intenzionale, ed è così che riusciamo a interpretare anche le scelte altrui da pochi indizi, spesso non verbali. Biologicamente parlando, sono gli stessi neuroni, quelli specchio, che si attivano quando immaginiamo un'azione, quando la eseguiamo e quando la vediamo eseguire. È grazie a questa capacità umana di simulare i comportamenti, di condividere le intenzioni e i loro legami con porzioni dello Sfondo effettivamente comune che possiamo evitare preamboli quando parliamo di Napoleone, perché siamo in grado di calarci immediatamente nei panni degli altri.

La soluzione, dunque, è intendere la realtà come qualcosa di molto più ampio, in cui il linguaggio, le pratiche e le semplificazioni stesse non sono esterni, ma da sempre aggiunte al dominio-realtà e comprese nel sistema. Le domande sul vero e sul falso andrebbero intese in senso non-assolutista, ma relativamente, in base ai motivi e al contesto in cui quei motivi trovano condivisione e senso. Pertanto, benché apparentemente simili nel contenuto, finiscono per essere tutte differenti nelle risposte e nelle condizioni di soddisfazione, a seconda degli scopi degli agenti. Insomma, l'esito è simile a quello di Wittgenstein nelle Ricerche Filosofiche<sup>35</sup>. La filosofia non è la scienza ed eliminare la prima perché non segue le regole della seconda è come dare il cartellino rosso secondo il regolamento del calcio a un atleta di kick-boxing che sta praticando quello sport sul ring.

La vera domanda a cui si vuole rispondere in questa sede non riguarda il fatto che il braccio che alzo sia di fatto alzato, ma come questo comunichi un'appartenenza fascista o come, in un altro contesto, lo stesso braccio indichi la richiesta di prendere parola in classe.

## Rappresentazioni

### Cultura

Le tecnologie risolvono problemi sociali, quindi non possono essere indipendenti dalla mente\corpo, dalle credenze, dalla morale e soprattutto da sentimenti e condizioni di felicità che le fanno andare a buon fine. Sono le applicazioni di leggi di laboratorio e, al tempo stesso, sono oggetti sociali non meno concreti.

Per semplificare gli usi, gli oggetti vengono strutturati con indicazioni all'usabilità: la mente\corpo viene iscritta nel design. Un tastierino di un videogame è fatto per mani di un certo tipo e contiene simboli su cui una cultura del gaming concorda e si tramanda: “Questo

<sup>34</sup> J. Searle, *Mente, Linguaggio e Società*, Scienze e idee, Raffaello Cortina Editore, Milano, 2000, p. 88.

<sup>35</sup> Wittgenstein L., *Ricerche Filosofiche*, tr. It. Einaudi, Torino, 1967.

simbolo X vale come SALTO nei videogame SONY<sup>36</sup>. Questo aspetto ci chiarisce ancora una volta quanto sia improprio parlare di oggetto o materiale o sociale. L'oggetto è, nel senso dato da Bruno Latour, ibrido<sup>37</sup>. Quando ci domandiamo del suo significato, a seconda dell'intenzione incorporata dalla domanda, siamo portati a premiare alle volte gli aspetti bruti, altre quelli culturali, altre ancora quelli individuali, senza che la risposta elimini le altre possibili, proprio perché l'oggetto non si riduce a nessuna componente, essendo ibrido, tra aspetti materiali e aspetti sociali.

Analogo discorso vale per il linguaggio, usato, tra le altre cose, per creare continui livelli di interpretazioni, e per fare della realtà un'esplosione di documenti, di istruzioni e di rappresentazioni, che vanno al di là della descrizione di ciò che è. Per un documento, il valore di vero o falso ha spesso senso politico, economico, emotivo: pensiamo alla falsità della Donazione di Costantino, essa non è come la falsità di una realtà composta da *omeomerie* (a parte il dispiacere personale di Anassagora, che tuttavia, ormai, lascia il tempo che trova). Ampliare il senso di reale, di verità come corrispondenza e di rappresentazione come metodo di organizzazione ci consente di comprendere anche i condizionali controfattuali<sup>38</sup>. Essi non vanno intesi sintatticamente o analizzati come asserzioni nella tavola di verità, altrimenti, come sappiamo, essi saranno tutti veri. Al contrario, il significato di tali proferimenti sta tutto nel sentimento che li accompagna. Rallegramento, sollievo, pentimento: questo il significato emotivo di quando si dice: "Se non avessi fatto così, sarebbe stato meglio\peggio". Da un punto di vista pragmatico, la funzione di tali sentimenti, con cui viene accompagnato il controfatto o con cui viene creata una contro-storia, è quella di diventare una guida per ripetere nuovamente quella condotta o per guardarsene ed evitarla in futuro. I condizionali controfattuali esistono perché, quando scegliamo, non valutiamo solo e soltanto sulla realtà del realista, ma grazie all'ego e all'incertezza che lo accompagna strutturalmente c'è un'esplosione di mondi possibili e credenze e possibilità da accogliere e scartare nella scelta. Dal momento in cui un'azione viene intrapresa, l'io cosciente ritorna sui mondi possibili non più possibili e sperimenta il sentimento associato ai controfatti. Questo è il miglior mondo possibile ma proprio grazie all'esistenza degli altri (mondi possibili).

I mondi possibili sono il risultato dell'aver fermato il flusso della vita, rappresentandolo in innumerevoli modi, dell'aver esercitato il dubbio, l'antirealismo in opposizione al realismo, sapendo usare la comprensione della differenza tra credenza e sapere, tra futuro e passato, tra necessità e contingenza. Ecco, allora, uno dei ruoli pratici del dubbio cartesiano, della filosofia, della rappresentazione. Insomma, la capacità rappresentativa degli esseri umani consente l'anticipazione e ogni ragionamento su mondi possibili. Si tratta di una guida per gestire la complessità.

<sup>36</sup> Gli atti costitutivi, secondo John Searle, sono un tipo di atti linguistici che non solo descrivono o rappresentano la realtà, ma la modificano o la costituiscono. Questi atti sono fondamentali per la creazione di istituzioni sociali e pratiche, come il matrimonio, la proprietà, gli accordi contrattuali, le promesse e così via. La forma delle regole costitutive, che, come si è detto, nel descrivere creano la realtà sociale, è la seguente: X conta come Y in C. (Cfr. Searl J. *La costruzione della realtà sociale* (1995))

<sup>37</sup> Latour, B. (2012, October 1). *We Have Never Been Modern*. Harvard University Press.

<sup>38</sup> Il condizionale controfattuale è una costruzione grammaticale che esprime una situazione ipotetica o immaginaria che si verifica nel passato e che non è effettivamente avvenuta. Viene utilizzato per descrivere eventi o situazioni che potrebbero essere accaduti, ma che in realtà non sono accaduti. Ad esempio, "Se avessi studiato di più, avrei superato l'esame" indica un'azione che non è stata effettivamente compiuta nel passato (lo studio) e le conseguenze ipotetiche di tale azione (il superamento dell'esame). Esso, analizzato secondo la tavola di verità del condizionale, sarebbe sempre vero, dato che l'antecedente ipotetico è falso "se avessi studiato", ma siccome non ho studiato, esprime un antecedente falso, che nella tavola di verità fa sì che la proposizione si sempre vera, qualunque sia il conseguente. Infatti, secondo la tavola di verità del condizionale,  $A \rightarrow B$  è F se  $A$  è Vero e  $B$  è Falso; mentre è vera in tutti gli altri casi. In Particolare quando  $A$  è vero,  $A \rightarrow B$  è necessariamente vera.

Lo Sfondo su cui agiamo è un insieme di layer, di presupposti all'agire e istituzioni create attraverso regole costitutive: si va da fatti più universali, specie-specifici, legati a qualunque forma di vita, fino a quelli individuali, basati su personalità, esperienze passate, abilità corporee personali, idioletto<sup>39</sup>. Tra i fatti bruti e le preferenze individuali ed embodied, esistono poi una serie di categorie sfumate da cui attingere, che possiamo chiamare culturali o di appartenenze a gruppi sovraindividuali. Esse non esistono in quanto tali, non c'è un quid netto che faccia esistere la cultura come categoria a sé. Ognuno di questi concetti si sfuma dentro a quel gruppo di soggetti che li metterà in pratica quando l'obiettivo lo richiede.

La cultura interviene a semplificare, flessibilmente, le scelte, ritagliando quel dominio altrimenti esageratamente complesso con strategie che probabilmente saranno utili in quel momento. Insomma, grazie al fatto di convivere in un gruppo e grazie al fatto di aver partecipato alle strategie auto-organizzative offerte dalle simulazioni, come descritto nei paragrafi precedenti, apprendiamo a semplificare la realtà in modo pre-attentivo, consentendo anticipazioni e interpretazioni delle scelte altrui che altrimenti richiederebbero troppo tempo e conoscenza. La cultura, diversamente da ciò che si potrebbe pensare, non aggiunge significato, ma aiuta gli esseri umani a selezionare informazioni da un mondo condiviso, reale, altrimenti ingestibile per le nostre risorse computazionali. È una riduzione automatica per gestire la ridondanza e superare il pensiero defettibile. La cultura può essere considerata una sorta di euristica (una fra le tante). Non appoggiamo, infatti, il relativismo forte, ma uno debole, in cui le fasi del cosiddetto ciclo cognitivo (percezione-azione) sono determinate dalla società, ma tale condizionamento non è fisso e inevitabile.

L'euristica culturale è una semplificazione che interviene in contesti specifici (sociali) ed è sempre aperta a cambiamenti da parte degli esseri umani nel corso delle loro azioni e goal in divenire. Qualsiasi influenza della cultura sulla cognizione può essere eliminata a favore di altre strategie apprese nei processi filogenetici o ontogenetici. Elisabetta Lalumera<sup>40</sup> ha dimostrato che anche il condizionamento del linguaggio sulla decisione dipende da come è stato impostato il task. Si chiama "flessibilità concettuale". Significa che un concetto è un insieme di strategie o procedure, alcune fornite dalla cultura, altre dal linguaggio, che l'individuo usa in modo flessibile, a seconda del tipo di compito.

Essere linguisticamente competenti non può prescindere dalla situazione, dagli interlocutori o dalla società che li ha dotati di strategie. La competenza linguistica (CL) come quella culturale (CC) è allora strettamente legata all'empatia e alla capacità di mettersi nei panni degli altri e di ragionare dalla loro prospettiva, per attuare la scelta migliore e per anticipare le scelte altrui, di modo da semplificare la complessità dei significati che potrebbero essere colti. La CC è legata alla rappresentazione, più o meno automatica, della realtà e al figurarsi mondi possibili su cui valutare le credenze più probabili.

Ecco allora il ruolo del dubbio, dell'allontanamento, delle simulazioni e dei controfatti: non sostituiscono la realtà, ma allenano la scelta flessibile dei piani ad agire più utili su una realtà strategicamente (e flessibilmente) ridotta.

## Tecnologie digitali: videogame, robot

<sup>39</sup> Lingua individuale, cioè la particolare varietà d'uso del sistema linguistico di una comunità che è propria di ogni singolo parlante (Cfr. idioletto - Treccani - <https://www.treccani.it/vocabolario/idioletto/>)

<sup>40</sup> Lalumera. E, 2013, Concetti, relativismo e strategie flessibili, Rivista italiana di filosofia del linguaggio, vol. 7, n. 3.

Come ogni tecnologia, sia videogiochi, sia robot, sia la realtà virtuale e aumentata sono tipologie di rappresentazioni<sup>41</sup>.

Come suggeriva Baudrillard<sup>42</sup>, c'è una sottile differenza tra *simulacra* e simulazioni. Nel secondo caso si tratta di rappresentazioni, di metafore che conservano un rapporto esplicito con l'originale, come può essere una mappa che imita il territorio, lo semplifica, mediando una relazione tra soggetto e realtà. La mappa, per permettere tale mediazione, non è solo una rappresentazione dell'ambiente, ma deve intrattenere un rapporto rappresentativo anche con l'essere umano, per il modo in cui rappresenta, impugna, significa. Quindi in un certo senso deve incorporare le specificità rappresentative dell'essere umano, i suoi percorsi simbolici e d'uso: in breve la sua mente e il suo corpo. L'interfaccia è sempre una rappresentazione dei due poli di collegamento. Il simulacrum, secondo Baudrillard, non intrattiene più rapporti con l'archetipo che rappresentava; perde il suo appoggio metaforico e diventa oggetto che rimanda a sé e non ad altro. Il quadro della Monnalisa non rimanda più alla modella ma a se stesso. Non è più principalmente un'imitazione, ma inizia una catena causale ed epistemica a sé. È, diremmo, agente deontico. Le tecnologie, spesso, attuano questo spostamento, da *representazioni di* a simboli in sé conclusi, che non mediano più alcuna realtà, né la nascono, ma sono la realtà, potendo diventare ponti per rappresentare altro, quindi l'originale per altri simulacri. Tuttavia, il fatto di essere simulazioni non scompare. Permane il rapporto con un originale di partenza, semplificato e trasdotto con la figura retorica della metafora. È proprio grazie a questo rapporto che avviene l'organizzazione dell'originale di partenza. Le scienze cognitive, sin dagli anni '70, hanno offerto prove incontrovertibili riguardo alla funzione cognitiva delle metafore. Queste figure retoriche sono impiegate dagli esseri umani per concettualizzare e semplificare un determinato dominio attraverso un altro, consentendo così di affrontare la complessità della realtà. Secondo Lakoff<sup>43</sup>, le metafore costituiscono il processo primario alla base del funzionamento della mente umana. Esse fungono da fondamenta per la categorizzazione, la cui essenza risiede nell'identificare ponti, termini medi e prototipi che collegano le istanze, facilitando così la fluidità concettuale. Il robot è un tipo particolare di metafora, nasce per diventare simulacro e per fare dell'essere umano con cui si relaziona una simulazione. Questo gioco di rimbalzi tra rappresentazione e rappresentato è circolo ma virtuoso. È come il gioco di rimbalzi di luce tra due specchi: La luce viene riflessa potenzialmente illimitate volte tra essi.

Il videogame simula un contesto, un ambiente di azioni, reazioni, oggetti e leggi fisiche credibili o incredibili, a seconda di quello che vogliamo ottenere; il robot, più del videogame, deve tener conto di quelle leggi fisiche, meccaniche, altrimenti non funzionerebbe, simulando sia l'agente sia come l'agente potrebbe simularsi il contesto. Nel videogame abbiamo il solo vincolo del computer, dell'hardware: esso deve rispondere alle sue regole fisiche e ai suoi limiti.

Il robot è differente rispetto alle altre rappresentazioni, metafore tecnologiche. Il robot ha il limite del mondo esterno e degli altri, con il loro terreno comune di presupposti con cui hanno dato senso al mondo stesso, nel quale un robot deve calarsi (*envelope*<sup>44</sup>). Ogni singola macchina robotica, a seconda del contesto, non simulato, in cui va a inserirsi, viene dotata di sensori e motori; nel farlo, prende ispirazione dalle strategie di risposta e quindi di sopravvivenza per quell'ambiente degli individui già selezionati dalla Natura da milioni di anni. Il robot per essere

<sup>41</sup> Saettone, L. (2023, June 26). *Videogame, robotica ed education: un rapporto sempre più stretto*. Agenda Digitale. <https://www.agendadigitale.eu/cultura-digitale/videogame-robotica-ed-education-un-rapporto-sempre-piu-stretto/>

<sup>42</sup> Baudrillard, J. (1981), *Simulacres et Simulation*, Éditions Galilée (French)

<sup>43</sup> G. Lakoff, Johnson M. (1998) *Metafora e vita quotidiana*, Bompiani (Milano).

<sup>44</sup> Per *envelope* si intende lo spazio tridimensionale in cui il robot deve agire. Non sempre si può strutturare uno spazio a misura di robot a cui gli esseri umani si adattano, ma si deve rendere il robot a misura di uno spazio già organizzato nel continuum con la mente estesa.



tale deve essere autonomo o semiautonomo quindi deve avere connessioni con l'ambiente, tramite sensori e motori e deve altresì avere una certa estetica che meglio lo fa rispondere ai vincoli ambientali, e anche a quelli sociali. In effetti se ci aspettiamo che un robot si interfacci con gli esseri umani (social robots) allora deve funzionare anche socialmente, incorporando caratteristiche anche culturali. Rappresentare la mente e il corpo degli esseri umani, interrogarci su come possa funzionare, assume un valore anche ingegneristico: per far funzionare meglio il robot ed adattarlo all'envelope degli esseri umani. Da questo punto di vista il robot diventa un simulacro: un agente causale ed epistemico.

Il robot deve rispondere alla necessità di adattamento al mondo, pertanto diventa come un'asserzione. Il suo descrivere, rappresentare, la mente e la realtà da esso intenzionata risponde a quella direzione di adattamento mente-a-mondo tipica delle credenze. Il robot è una parola, esprime la forza di una credenza vera o falsa se si adatta ai mondi possibili (su cui gli atteggiamenti epistemici sono veri o falsi). Pertanto includerli nell'architettura cognitiva potrebbe essere ingegneristicamente vincente<sup>45</sup>. Per operare efficacemente in ambienti sociali e culturali complessi, il robot deve anche essere in grado di simulare la mente altrui, adottando una struttura cognitiva ibrida che integri moduli per la percezione, la memoria, l'affettività e il ragionamento. L'architettura cognitiva del robot deve essere in grado di adattarsi dinamicamente al contesto sociale e culturale, utilizzando strategie di selezione delle credenze plausibili e simulando una serie di mondi possibili coerenti con la situazione attuale. Questo approccio permette al robot di sviluppare e aggiornare continuamente le proprie conoscenze, personalizzando le proprie azioni in base alle richieste ambientali e alle interazioni con gli esseri umani. Inoltre, il robot deve essere in grado di implementare euristiche umane, come l'empatia e l'introspezione, per anticipare e adattarsi rapidamente alle esigenze e alle preferenze degli individui con cui interagisce.

Come abbiamo anticipato fu Saul Kripke<sup>46</sup> ad aver introdotto una semantica per i contesti modali, specificamente per le situazioni in cui sono coinvolti verbi di atteggiamento proposizionale come credenze e conoscenze. Le asserzioni come 'Lorenza sa che  $\alpha$ ' o 'Lorenza crede che  $\alpha$ ' non seguono le tavole di verità della logica proposizionale tradizionale, poiché la verità o falsità dipende dalla relazione tra le credenze del soggetto e i fatti del mondo attuale. Questo approccio modale indica che credere o sapere qualcosa significa immaginare mondi possibili, dove la verità è valutata su situazioni possibili determinate in base a una relazione di 'compatibilità' tra le credenze del soggetto, comprese quelle che avrebbero potuto e potrebbero essere, e i fatti del mondo attuale. Non si può 'sapere' che il "Lucca Comics" è una sagra delle salsicce perché è falso; si può, tuttavia, 'credere' che sia una sagra delle salsicce, dato che l'operatore 'credere', in 'credere che P', è meno forte. Per definire quello che si crede e che si sa di credere, possiamo impiegare diverse forme di funzioni di accessibilità tra mondi possibili. La più completa delle Relazioni tra mondi è S5, ed include l'introspezione negativa: cioè il fatto eventualmente di sapere di non sapere qualcosa. Solo a partire da questa consapevolezza di ignoranza una persona è spinta a cercare le informazioni mancanti. In breve, il modello rappresentativo dei mondi possibili permette dinamicamente di aggiornarsi sulla base di credenze probabilistiche offerte da motivazioni e contesto, grazie alla relazione S5 che simula introspezione, attenzione selettiva ed empatia. Tenta una partenza sulla base della più probabile delle credenze, la più frequente nei mondi possibili che riusciamo a immaginare. Infatti gli stati alternativi che si pongono alla nostra immaginazione non rappresentano situazioni casuali, ma le formule presenti in esse sono distribuite sulla base di calcoli probabilistici. L'allocazione di

<sup>45</sup> L. Saettone, R. Fedriga, E. Micheli. (2023), Linguistic and Cultural Competencies in Dynamic Possible Worlds, ROMAN conference, Robotics Workshop, Busan.

<sup>46</sup> Kripke S, op cit.

tali credenze dipende, insomma, da ciò che è ragionevole credere, pur mantenendo la consapevolezza (S5) che una credenza non è sapere (necessariamente vera), ma è contingente: è possibile che si venga smentiti. Infatti, si assume che il sistema venga personalizzato, alla luce di eventuali rettifiche da parte dei fatti del mondo attuale, e quindi la credenza diventa sapere. Si tratta di un modello nel quale si è tentata una prima risposta in base alla più probabile delle alternative, a partire da informazioni anche culturali, pre-giudizi, cioè pre-comprensioni. Il sistema è dinamico per l'aggiornamento delle credenze circa le informazioni disponibili e quelle da aggiornare o corroborare. Questo approccio, chiamato anche socratico, permette di attualizzare azioni diverse per modificare il mondo, trasformare la conoscenza degli altri e perfezionare la propria, tenendo conto della dinamicità dei mondi possibili e dell'indeterminatezza del contesto. Questo modello consente, insomma, di evitare bias culturali fissi, il pregiudizio è una euristica. Infatti ci consente di offrire una prima risposta, senza rimanere inermi nell'attendere più informazioni che potrebbero non giungere, attingendo dalla macro-categoria sociale, come principio di economia. Grazie a questo bacino sovra-individuale si possono tentare interpretazioni, ma grazie al sistema S5 e alla possibilità dinamica di aggiornare i mondi possibili con nuove informazioni e nuovi pesi associati alle credenze, corroborate o smentite, si possono interpretare situazioni in modo flessibile e sempre adattivo, riflettendo la natura complessa e probabilistica delle interazioni umane. Quello che sembra ridondante, un'esplosione di mondi aggiuntivi oltre a quello attuale, in realtà è una semplificazione della complessità, un modo per simulare attenzione, rappresentazioni e ragionamenti abduttivi, in mancanza di certezza e tempo.

Abbiamo detto, quindi, che la cultura ci fornisce euristiche già testate dal gruppo per semplificare l'elaborazione delle informazioni nelle stesse condizioni. Il tipo di compito è come uno stimolo che ci fa richiamare i processi correlati, facendoci utilizzare le migliori strategie per risolvere ogni compito. Ogni lingua e ogni cultura hanno sicuramente peculiarità specifiche con cui interpretano il mondo. Dotano gli individui di piani d'azione, ma nessuno di essi è inevitabile: non siamo burattini nelle mani del linguaggio e della cultura, come se indossassimo occhiali con i quali potremmo vedere il mondo solo attraverso un tipo di lente colorata. Invece, ciò che chiamiamo 'categoria' non è altro che un database di rappresentazioni e procedure apprese in modi diversi, ognuna richiamabile in base alla sua utilità e rinforzo.

Nei suoi studi, Elisabetta Lalumera<sup>47</sup> ha dimostrato che il condizionamento del linguaggio dipende da come viene impostato il task. I risultati del determinismo linguistico derivano sempre dal contesto e dal tipo di compito, quindi cambiando le richieste, possiamo anche annullare gli effetti cognitivi del linguaggio. Quando categorizziamo, quando svolgiamo compiti di risoluzione dei problemi e facciamo esercizi di inferenza e riconoscimento, utilizziamo gli indizi che, ad esempio, la nostra lingua ci fornisce. È un principio di risparmio economico. Siamo computer limitati in termini di potenza e calcolo, quindi naturalmente sfruttiamo le euristiche che già conosciamo.

Un concetto non è qualcosa di fisso, ma un insieme di strategie e procedure che l'individuo ha sempre a disposizione e può utilizzare in modo flessibile, a seconda del tipo di compito. Possedere un sistema concettuale è come avere una rete di competenze con uno scopo comune. Si tratta di interazioni con il mondo, che possono realizzarsi in modi diversi. I concetti sono flessibili, il che significa che non sono rappresentazioni mentali unite da un'essenza fissa; invece, sono funzioni la cui realizzabilità è molto ampia. Insomma, interrogarci su come possa essere simulata tale caratteristica della mente e dell'intenzionalità consente da un lato di diventare più organizzati nel farlo, acquisendo un habitus, dall'altro diventa una semplificazione per il robot o per l'avatar dotarlo di un'architettura flessibile, capace di

---

<sup>47</sup> Lalumera E., op. cit.

ragionare su mondi possibili e probabilità e di adeguare i piani in base a pesi diversi associati a motivi diversi, cambiando quegli stessi pesi alla luce di feedback esterni, e quindi permettere la personalizzazione<sup>48</sup>.

## Tecnologie digitali: VR, AR

Diversi ancora i casi della Realtà Aumentata (AR) e della Realtà Virtuale (VR); si tratta di simulazioni che pongono in essere il ruolo epistemologico dell'essere umano, la sua presenza in due tipologie di reale.

L'AR è una tecnologia che mostra come avviene la formazione di layer documentali sui singoli oggetti<sup>49</sup>. L'AR si sovrappone a ciò che ci circonda, a una realtà che resta sempre ben visibile da noi e della telecamera del *device*; è necessario che il Dominio di partenza resti percepibile come base su cui il software, dopo il riconoscimento, proietterà un contenuto digitale sovrapposto, la sua metafora.

Diverso ancora è il caso della realtà virtuale (VR), la quale tende a *darsi come* unica realtà. Il visore di VR è come un paraocchi, studiato per isolarci a livello visivo da ciò che ci circonda. Come quello degli equini, esso ha la funzione di non far distrarre l'animale, portandolo a concentrarsi solo sul rettangolo di sentiero immediatamente davanti a lui. Analogamente il visore ha lo scopo di distogliere da ogni stimolo visivo esterno, dando una apparenza di partecipazione a una realtà altra. È una sorta di caverna platonica, ma luminosa, di pixel. Ben rappresenta la prova che la coscienza non sia riproducibile in una capsula di Petri, secondo stimolazione appropriata.<sup>50</sup> Infatti, quando siamo calati nel contesto virtuale, la separazione dalla realtà estesa, che il visore cerca di celare, è immediatamente smascherabile. Il nostro corpo è collocato e aperto tramite gli altri sensi nella realtà parzialmente obliterata; pertanto, essa è sempre esperibile in ogni istante dal soggetto. La vista da sola non è sufficiente a dare percezione del reale, ma agisce in concerto ad altre modalità sensoriali e alle possibilità su di esse. Nel mondo virtuale anche le intenzioni e il libero arbitrio sono limitati: le azioni e la creatività sono quelle incluse nel gioco, eseguibili entro i limiti della macchina stessa: insomma i gradi di libertà concessi sono sempre uguali o inferiori a quelli del sistema e rispetto a quelli a cui ci ha abituati il Dominio del Reale sono indubbiamente irrisori, condannandoci a una sorta di insonnia inevitabile<sup>51</sup>. Infatti, ogni volta che si prova ad agire al di là dei vincoli di default e ogni volta che i sensi non si allineano (quello che si vede non è quello che si tocca, sente, annusa) si genera una dissonanza cognitiva che porta al risveglio.

A differenza della Realtà Virtuale VR che tende a spacciarsi per unica realtà, l'AR si sovrappone a ciò che ci circonda, a una realtà che resta sempre ben visibile da noi e della telecamera del *device*. L'AR è molto più affine alla teoria degli effetti di Gadamer<sup>52</sup>: entrambe aggiungono interpretazioni al testo originale, il quale si aumenta di sensi (taciti o espliciti), prendendo vie inattese rispetto all'intenzione primigenia. La realtà aumentata può aggiungere

---

<sup>48</sup> L. Saettone et al, Cultural Competence through Dynamic Epistemic Logic: a proposal for robotic implementation, IRIM conference, 2023, Rome.

<sup>49</sup> Per "Layer documentali" intendo la catena di regole costitutive che porta a costruire la realtà sociale, in cui dopo un primo momento in cui avevamo nella sostituzione metaforica una partenza materiale, poi pian piano la abbandoniamo, utilizzando simboli sociali come se fossero reali, per costruire ulteriori domini altrettanto sociali. Si viene a creare una sovrapposizione stratificata di istituzioni e di istituzioni di istituzioni, nella stessa maniera in cui un robot posto di fronte a un essere umano gioca all'infinito a ping pong con le metafore.

<sup>50</sup> Noe A, 2010, Perché non siamo il nostro cervello, Raffaello Cortina Editore, Milano.

<sup>51</sup> Sisini F. Sisini V, art cit.

<sup>52</sup> Grondin, J. (2003, January 1). *Hans-Georg Gadamer*. New Haven : Yale University Press.

un contesto di lettura all'oggetto-base, spostandone completamente il senso, o chiarendone la leggibilità come un'*affordance* con legenda.

Per altri versi l'AR somiglia alla pratica memica: c'è un originale ma aperto alle declinazioni e applicazioni degli altri soggetti, che ne replicano innumerevoli volte la struttura. C'è una struttura-base (il fatto) a cui si aggiungono layer interpretativi. Anche il meme consta di una base di partenza, che non è il riferimento, ma il suo senso<sup>53</sup>. Il riferimento, il contenuto, viene dato successivamente dagli agenti che inseriscono la figura del meme in infiniti contesti. Il meme-struttura è un atto perlocutorio<sup>54</sup>: Bernie Sanders (Figura 1) con le muffole è come una figura retorica analizzata nella pragmatica; è l'equivalente di un romano medio che dice "stic\*\*\*" quando Cesare oltrepassa il Rubicone, ma lui deve pensare a come arrivare a fine giornata. Bernie Sanders con le muffole, nel diventare meme, viene ritagliato da Capitol Hill e si inserisce in qualunque situazione altra, con il senso implicito di schernire la serietà di quel dato contesto. Quando una struttura memica è declinata secondo contesti diversi, il contenuto cambia ma non gli effetti sui parlanti o la forza e le intenzioni primigenie (una promessa è sempre una promessa come Bernie Sanders con i guantoni significa disinteresse cinico per il contenuto a cui si applica).

Con la Realtà Aumentata la figura-marker che la attiva è il riferimento, mentre le informazioni digitali sono il non-detto della pragmatica: esiste un oggetto, che, se inquadrato, si arricchisce di nuovi elementi, quelli che nella comunicazione sono dati per scontato: le implicature, che pertanto mutano in espliciture. Le implicature si riferiscono a ciò che viene implicitamente suggerito o implicato da un'asserzione. Tutto ciò che viene lasciato sott'inteso, dipende da un terreno comune socialmente costruito, garantito anche dall'accesso a quei mondi possibili descritti in precedenza. Dire che "Roberta beve" lascia sott'inteso "beve alcol", a meno che non venga reso esplicito "Roberta beve... bibite". Una Realtà Aumentata è come se volesse esplicitare le *affordances* della Realtà, aggiungendo, in digitale, layer studiati appositamente per creare compresenza esplicite di ipertesti.

---

<sup>53</sup> Secondo Frege, ogni espressione linguistica ha un senso, che rappresenta il modo in cui è compreso il significato che incorpora un'espressione, e un riferimento, che indica l'oggetto o concetto a cui si riferisce nell'universo del discorso. Discepolo di Platone e Maestro di Alessandro Magno hanno lo stesso riferimento: Aristotele; ma hanno senso diverso: nel primo caso per conoscere il senso di quella descrizione devo guardare a Platone e a chi partecipava nell'accademia, nel secondo caso devo consultare altri documenti storici che si riferiscono al Macedone che conquistò fino all'India. (Cfr. rege, G. (1892). Über Sinn und Bedeutung. Zeitschrift für Philosophie und philosophische Kritik, 100, 25-50. Trad. it. parziale in Frege (1965) e trad. it. integrale in Bonomi (1973))

<sup>54</sup> gli atti perlocutori sono azioni che vengono compiute attraverso il linguaggio e che influenzano il comportamento o le percezioni degli altri. (Cfr Searle, J. R. (2009, January 1). Atti linguistici. Saggi di filosofia del linguaggio.

[http://books.google.it/books?id=kLfpNwAACAAJ&dq=searle+atti&hl=&cd=1&source=gbs\\_api](http://books.google.it/books?id=kLfpNwAACAAJ&dq=searle+atti&hl=&cd=1&source=gbs_api))



Figura 1 Bernie Sanders

La Realtà Aumentata mostra su schermo le aggiunte al testo di partenza, aggiunte ipertestuali che aumentano o cambiano il significato. Essa ci chiarisce che a. La base è reale e si può attenzionare in modo flessibile b. Il significato, il meta-dato sono reali essi stessi e flessibili. Insomma, la Realtà Aumentata va a rappresentare esattamente il modo in cui si creano istituzioni e fatti sociali, documenti, quindi le rappresentazioni e come intenzionalmente la coscienza performa i suoi molteplici legami con il reale.

## Conclusioni

Epoché, domande filosofiche controintuitive e rappresentazioni (comprese le tecnologie) organizzano i nostri legami con la realtà, la quale resta indipendente e disponibile a nuove negoziazioni. Insomma, l'antirealismo non va inteso quale sostituto della realtà, la quale, anzi, risulta oltremodo indispensabile al fine di produrre la sua negazione, come resta indispensabile per "triggerare" la Realtà Aumentata o per collocarvi i robot, tecnologie che rendono visibili le pratiche sommerse con cui gli esseri umani costituiscono i legami intensionali con le cose, accelerando la comprensione senza necessità di fare preamboli.

In questo articolo abbiamo preferito evitare categorie come mente-corpo, mostrando come invece la realtà emerga esattamente nella triangolazione di menti, non dissimili, e mondo in comune. Nell'agire quotidiano le attese vengono ricalibrate costantemente, co-creando uno sfondo effettivamente comune in cui layers di vario tipo e strategie concettuali molteplici si innestano semplificando azioni ed empatia. In questo processo la filosofia, l'arte, le tecnologie diventano strumenti, metodi, per fermare temporaneamente la coazione a ripetere e riorganizzare le attività di base che rappresentano, aggiornando o confermando i presupposti dell'intenzionalità collettiva.

## Bibliografia

- Austin, John Langshaw. *Sense and sensibilia* (1959), Oxford: Oxford UP, 1964 ; Putnam, Hilary. *Mente, corpo, mondo*. Il Mulino, 2003
- Austin j. l. (1962). *how to do things with words*. oxford university press
- Baudrillard, J. (1981), *Simulacres et Simulation*, Éditions Galilée (French)
- Andy Clark e David J. Chalmers, *The Extended Mind*, in *Analysis*, vol. 58, n. 1, gennaio 1998, pp. 7-9, DOI:10.2307/3328150
- Frege, G. (1892). *Über Sinn und Bedeutung*. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25-50. Trad. it. parziale in Frege (1965) e trad. it. integrale in Bonomi (1973)
- Frixione M., Iaquinto S., Vignolo M., (2016) *Introduzione alle logiche modali*, Laterza, Roma Bari.
- J. Gibson. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979.
- Grondin, J. (2003, January 1). *Hans-Georg Gadamer*. New Haven : Yale University Press.
- Kripke, S 1980 *Naming and Necessity* Cambridge Harvard University Press
- Latour, B. (2012, October 1). *We Have Never Been Modern*. Harvard University Press.
- G. Lakoff, Johnson M. (1998) *Metafora e vita quotidiana*, Bompiani (Milano).
- Lalumera, E, 2013, *Concetti, relativismo e strategie flessibili*, *Rivista italiana di filosofia del linguaggio*, vol. 7, n. 3. Noè. *Strange Tools: Art and Human Nature*. First edition. New York, Hill and Wang, a division of Farrar, Straus and Giroux, 2015.
- Noe A., *Perché non siamo il nostro cervello*, 2010, Raffaello Cortina Editore, Milano
- Popper, K.R., *The poverty of historicism*, London 1957 (tr. it.: *Miseria dello storicismo*, Milano 1975
- Realism*, in *Stanford Encyclopedia of Philosophy*, 2019, December 13. <https://plato.stanford.edu/entries/realism/>
- L. Saettone, R. Fedriga, E. Micheli, (2023), *Linguistic and Cultural Competencies in Dynamic Possible Worlds*, RobOntics 2023 Ontologies for Autonomous Robotics 2023 Proceedings of the Workshop on Ontologies for Autonomous Robotics co-located with The 32nd IEEE International Conference on Robot and Human Interactive Communication (IEEE RO-MAN 2023) Busan, S. Korea, August 28, 2023.
- Saettone, L. (2023, June 26). *Videogame, robotica ed education: un rapporto sempre più stretto*. Agenda Digitale. <https://www.agendadigitale.eu/cultura-digitale/videogame-robotica-ed-education-un-rapporto-sempre-piu-stretto/>
- L. Saettone, R. Fedriga, E. Micheli. (2023), *Linguistic and Cultural Competencies in Dynamic Possible Worlds*, ROMAN conference, Robotics Workshop, Busan.
- L. Saettone et al, *Cultural Competence through Dynamic Epistemic Logic: a proposal for robotic implementation*, IRIM conference, 2023, Rome.
- Searle J. *La costruzione della realtà sociale* (1995)
- J. Searle, *Mente, Linguaggio e Società*, Scienze e idee, Raffaello Cortina Editore, Milano, 2000.
- Searle, J. R. (2009, January 1). *Atti linguistici*. *Saggi di filosofia del linguaggio*. [http://books.google.ie/books?id=kLfpNwAACAQJ&dq=searle+atti&hl=&cd=1&source=gbs\\_api](http://books.google.ie/books?id=kLfpNwAACAQJ&dq=searle+atti&hl=&cd=1&source=gbs_api)
- Sisini, Francesco & Sisini, Valentina & Sisini, Laura. (2023). *Limiti teorici della Virtual Reality*. 10.13140/RG.2.2.27449.65126.
- Tomasello, M. (2009, June 30). *Constructing a Language*. Harvard University Press.
- Wittgenstein L., *Ricerche Filosofiche*, tr. It. Einaudi, Torino, 1967.

# Perché l'AI non potrà mai replicare l'essenza dell'intelligenza umana

**Intelligenza umana e artificiale sono e rimarranno sempre diverse. L'AI potrà anche arricchirsi di una conoscenza enorme ma è e resterà sempre sprovvista di un corpo e della componente affettiva, sempre in gioco nell'acquisizione di nozioni e nel potenziamento delle nostre capacità**

Di Roberto Pozzetti, Psicoanalista, Professore a contratto LUDeS Campus Lugano, Professore a contratto Università dell'Insubria, autore del libro 'Bucare lo schermo. Psicoanalisi e oggetti digitali', già referente per la provincia di Como dell'Ordine degli Psicologi della Lombardia

La diffusione degli apparecchi di intelligenza artificiale suscita una serie di reazioni cariche di preoccupazione e inquietudine, come avviene dinanzi a ogni nuova scoperta che si annuncia rivoluzionaria sino a promettere una radicale discontinuità. Capi di Stato ed esperti di tutto il mondo, selezionati nell'ambito della comunità di ricerca, del mondo accademico e della società civile si riuniscono per approntare sistemi di sicurezza in questa dimensione e sottoscrivono un codice etico a tale proposito. Siamo, però, sicuri che l'intelligenza artificiale sarebbe effettivamente in grado di soppiantare quella umana?

## Cosa intendiamo per Intelligenza Artificiale

Non è semplicissimo circoscrivere cosa vada situato nel novero dell'intelligenza artificiale, termine coniato, alla metà degli anni Cinquanta, dall'informatico statunitense John McCarthy, vincitore del premio Turing nel 1971. D'altronde, non è neppure facile delimitare il concetto di dispositivo digitale; per esempio, come sottolinea il filosofo Luciano Floridi, quando siamo alla guida della nostra utilitaria ci troviamo offline oppure online? Ci sembra di essere offline in quanto concentrati a osservare la strada dinanzi a noi ma ci avvaliamo della digitalizzazione utilizzando il navigatore o Google Maps e parlando al telefono grazie al Bluetooth.

Formati immediati e ormai comuni quanto all'intelligenza artificiale sono ChatGPT, inventata da Sam Altman, cui si è aggiunto, più di recente, Seenapse che suggerisce idee in modo creativo, fra l'altro per rendere più convincente un programma pubblicitario, per comunicare un proprio messaggio in modo accattivante. Del resto, quasi tutti utilizziamo frequentemente altre forme di Intelligenza Artificiale come l'assistente virtuale domestico e l'assistente vocale dei nostri smartphone. Le tecniche di Intelligenza Artificiale si basano sull'opportunità offerta dall'attuale informatica di scrivere programmi che vanno a replicare il funzionamento della nostra mente. Attualmente sappiamo inventare sistemi in grado di apprendere dall'esperienza, per citare il titolo di un famoso libro dello psicoanalista inglese Bion. Quella sequenza di operazioni elementari e non ambigue che viene detta algoritmo non è più applicata in modo rigido, come avveniva con i computer di precedenti epoche. In questo modo abbiamo congegni digitali che non si limitano a osservare la realtà e a intervenire su di essa in base a come sono programmati ma, in un certo qual modo, riescono a interpretarla come è in grado di fare una persona.

Si diffondono e si diffonderanno ancor più in futuro strumenti in grado di rimpiazzare gli esseri umani nello svolgimento di attività impiegate su conoscenze teoriche e capacità logiche in precedenza specifiche dell'umanità.

## Le quattro rivoluzioni che spodestano l'io

Luciano Floridi ha coniato il termine *onlife* per descrivere le nostre giornate nelle quali valichiamo sovente la differenza fra vita *online* e vita *offline*. Egli scrive a proposito della quarta rivoluzione, successiva alle tre già messe in risalto a suo tempo da Freud e che hanno suscitato “umiliazioni” negli esseri umani.

La rivoluzione cosmologica copernicana ci spodesta dall'idea della Terra come centro dell'universo scoprendo che al centro del sistema solare vi è il Sole.

La sovversione biologica darwiniana mostra come l'uomo sia frutto dell'evoluzione animale anziché di una creazione che ci renderebbe a immagine di Dio.

La terza rivoluzione, quella relativa alla scoperta del desiderio inconscio compiuta da Freud, ribalta il *cogito ergo sum* cartesiano. Crediamo di pensare ma siamo, in effetti, parlati dall'inconscio. Pensiamo di volere qualcosa ma il desiderio inconscio che emerge nei nostri sogni ci conduce a sovvertire tale pregiudizio. La rivoluzione freudiana dimostra che l'inconscio orienta la nostra mente e la nostra esistenza: “l'io non è padrone in casa propria” – questo il celebre aforisma di Freud. È l'inconscio, strutturato come un linguaggio, l'unico padrone.

La quarta rivoluzione, indotta dal digitale, sovverte l'ingenua illusione dell'uomo come essere intelligente senza eguali. La macchina di Turing, detta anche computer, ci supera nella capacità di processare informazioni: non siamo più gli indiscussi proprietari della capacità di agire e operare in modo intelligente. In effetti, già il Lacan degli anni Cinquanta si era interessato alla cibernetica, inventata pochi anni prima dall'ingegnere statunitense Norbert Wiener: aveva notato come questa lasciasse intatti i presupposti della pratica della psicoanalisi in quanto sono strutturati come un linguaggio sia la cibernetica, con l'alternarsi delle cifre 0 e 1, sia la psicoanalisi imperniata sul linguaggio quale condizione dell'inconscio.

## A proposito del concetto di intelligenza

Questo dibattito sui rischi correlati all'espansione dell'intelligenza artificiale che potrebbe interferire con le attività umane, analogamente a quanto avviene con il supercomputer HAL 9000 nel celebre film di Stanley Kubrick “2001 Odissea nello spazio”, sembra dare per scontato che vi sia uniformità di punti di vista circa il concetto di intelligenza. In effetti, in psicologia, l'intelligenza viene considerata in diverse forme: si parla di intelligenza linguistica, logico-matematica, emotiva, musicale, corporeo-cinestesica, eccetera. L'intelligenza viene anche misurata e in questo ebbe un ruolo cruciale lo psicologo Alfred Binet che inventò dei metodi per calcolarla. Viene misurata con test come la WAIS (Wechsler Adult Intelligent Scale) e la WISC (Wechsler Intelligence Scale for Children) strutturati in subtest che appartengono in parte al campo dell'intelligenza verbale e per la restante parte a quella dell'intelligenza volta alla performance oppure con quello delle matrici di Raven. Attribuendo un valore di 100 ai risultati medi conseguiti da un soggetto in base alla sua età cronologica, si considera che vi sia un significativo deficit dinanzi a esiti inferiori al valore di 70 e una plusdotazione intellettiva in caso di risultati che superano i 130. Nel range intermedio, fra questi valori di soglia, si ritiene di aver a che fare con un'intelligenza tutto sommato nella norma.



In estrema sintesi, a livello etimologico, intelligenza è un termine che deriva dal latino *intus legere*: dunque l'intelligenza concerne un complesso di facoltà mentali relative a un leggere dentro, a una lettura interiore, a un cogliere qualcosa approfonditamente. Si teme che il ricorso all'intelligenza artificiale vada a soppiantare quella umana e determini, dunque, un incremento della disoccupazione; in effetti, senza distruggere i macchinari come proponevano invece i luddisti, tutti gli avanzamenti del progresso hanno determinato il sorgere di nuove professionalità e di ulteriori posti di lavoro con una riqualificazione delle competenze richieste.

Di fatto, il concetto di intelligenza stesso è molto opinabile; il riuscire o meno a svolgere degli specifici compiti si dimostra maggiormente relativo all'attuazione di un metodo di studio oppure di lavoro appropriato, allo sviluppo di strategie efficaci, all'impegno nel potenziamento delle proprie abilità, dei propri skills anziché a qualcosa di innato e imm modificabile. A proposito della presunta intelligenza e delle tesi relative a una demenza digitale, proposto dal neuropsichiatra e professore universitario tedesco Manfred Spitzer, abbiamo già scritto un articolo pubblicato qui in data 8 marzo 2023 nel quale sosteniamo a proposito dei nativi digitali che, anziché trattarsi di soggetti demenziali vale a dire con lacune mentali, riscontriamo degli altri tipi di mente: sono meno capaci sul piano mnemonico ma sviluppano maggiormente competenze logiche.

## L'intelligenza umana non potrà mai venire sostituita da quella artificiale

Le due forme di intelligenza, di lettura interiore, sono e rimarranno sempre diverse. È questa la nostra tesi fondamentale. Matteo Pasquinelli, professore di filosofia della scienza a Cà Foscari a Venezia, sottolinea come questi timori relativi all'intelligenza artificiale siano dovuti anche a delle forme di fallacia logica; una di queste sta nell'attribuirle ingenue analogie con quella umana soprattutto in termini di una volontà minacciosa e persecutoria. Pensiamo alla consuetudine di costruire dei robot a immagine e somiglianza degli uomini: questo ci porta ad attribuire alle macchine un funzionamento simile a quello umano che viene tuttavia smentito dai fatti.

### Il ruolo del corpo nell'apprendimento e nello sviluppo dell'intelligenza

Quello che appare fondamentale a proposito della differenza fra intelligenza artificiale e intelligenza umana sta nel ruolo del corpo nell'apprendimento e nello sviluppo dell'intelligenza, come sottolinea il noto neurologo e psicologo di origine portoghese Antonio Damasio. Un bambino sviluppa la propria intelligenza imparando anche e soprattutto dalle proprie esperienze corporee: impara a tenersi a debita distanza da fonti di calore elevato dopo essersi ustionato una mano, apprende la capacità di stare in equilibrio dopo aver giocato al girotondo, apprende come trarre piacere dopo aver esplorato le proprie parti del corpo traendone piacere. Qualunque dispositivo tecnologico, per quanto sia dotato di Intelligenza Artificiale, rimarrà comunque privo di un corpo e del godimento intrinseco al corpo; secondo Lacan, infatti, il corpo umano è caratterizzato dal godimento che può avvenire in svariate forme singolari. La studiosa inglese Isabel Millar dell'Università del Kent ha posto al centro di un suo recente libro, a proposito dell'intelligenza artificiale, il seguente provocatorio interrogativo: "Gode?". Le macchine sono in grado di ampliare le proprie conoscenze ben al di là di quelle umane ma rimangono sprovviste dell'esperienza del godimento che si ritrova soltanto nell'essere parlante.

### La componente affettiva

Molto rilevante è inoltre la componente affettiva, sempre in gioco nell'acquisizione di nozioni e nel potenziamento delle proprie capacità. Lo si percepisce francamente quando l'intelligenza si articola con il campo dell'amore: dunque, per esempio, un allievo in contesti scolastici ottiene riscontri più validi quando si trova bene con un docente che gli ispira simpatia e che riesce ad appassionarlo alla materia da lui insegnata mentre avrà un calo del rendimento scolastico se giungerà a instaurarsi un clima di tensione o di pressione con quell'insegnante che lo porterà a demotivarsi. In quest'ultimo caso, si verifica un'inibizione delle potenzialità intellettuali per ragioni di ordine affettivo.

Un dispositivo dotato di Intelligenza Artificiale è tale perché viene inventato e programmato da individui umani, anche a livello affettivo. Celebre in questi termini un'altra raffigurazione cinematografica: quella dei replicanti del film *Blade Runner* di Ridley Scott, tratto dal libro "Il cacciatore di androidi" di Philip Dick, i quali sembrano acquisire emozioni e affetti umani fino al termine della loro esistenza, con addirittura esperienze e potenzialità in più rispetto al genere umano: "Io ne ho viste cose che voi umani non potreste immaginarvi. Navi da combattimento in fiamme al largo dei bastioni di Orione. E ho visto i raggi B balenare nel buio vicino alle porte di Tannhauser. E tutti quei momenti andranno perduti nel tempo, come lacrime nella pioggia. È tempo di morire".

## Conclusioni

Noi sosteniamo che le due forme di intelligenza sono e rimarranno sempre diverse. L'intelligenza artificiale può arricchirsi di una conoscenza enorme ma è e resterà sempre sprovvista di un corpo, rimarrà priva di quelle forme di godimento che contraddistinguono il corpo umano. L'idea di un pericolo indotto dall'intelligenza artificiale è soprattutto una nostra fantasia.

## Bibliografia

Bion W. R., *Apprendere dall'esperienza*, Armando Armando Editore, Roma, 1983.

Floridi L., *La quarta rivoluzione*, Raffaello Cortina, Milano, 2017.

Lacan J., *Il seminario. Libro II. L'io nella teoria di Freud e nella tecnica della psicoanalisi*, Einaudi, Torino, 1991.

Millar I., *The psychoanalysis of Artificial Intelligence*, The Palgrave Lacan Series, 2021.

Pasquinelli M., *The eye of the master: a social history of Artificial Intelligence*, Verso, Londra, 2023.

# L'AI può aiutarci a “leggere” i pensieri? Progressi concreti e dubbi etici

Neuralink, intelligenza artificiale e risonanza magnetica funzionale sono solo alcune delle tecnologie usate per "leggere" i pensieri. Ma di fianco a queste soluzioni emergono anche complessi interrogativi sulla sicurezza e l'integrità del singolo nel panorama delle neuroscienze

Di Tommaso Ciulli, Psicologo

La letteratura ed il cinema ci hanno regalato nel tempo centinaia di racconti e film di fantascienza o supereroi dove personaggi di fantasia, riescono con qualche tecnologia o potere a leggere la mente delle persone. Ebbene, l'antico sogno dell'uomo di svelare i segreti nascosti nella profondità dei pensieri sta lentamente ma inesorabilmente diventando una realtà tangibile, grazie alle nuove frontiere della neurotecnologia.

Tuttavia, tra progressi tecnologici e questioni etiche, le sfide da affrontare sono molteplici: l'intelligenza artificiale e la risonanza magnetica funzionale rappresentano solo alcune delle soluzioni innovative adottate per "leggere" i pensieri umani, ma portano con sé interrogativi complessi sulla sicurezza e sull'integrità del singolo nel vasto panorama delle neuroscienze emergenti.

## Il desiderio umano di leggere la mente

Da sempre è un desiderio dell'essere umano avere la possibilità di leggere la mente dell'altro. Potremmo dire che questo bisogno sia legato alla naturale tendenza dell'uomo a volere avere relazioni quanto meno stabili e positive, ma anche dalla spinta della nostra mente a voler sapere e spiegare cosa accade al di fuori di essa, anche per influenzarne l'andamento.

Nella nostra evoluzione abbiamo sviluppato alcune abilità adibite proprio alla “lettura” della mente dell'altro. In questo caso parliamo di capacità metacognitive [1,2] che ci aiutano nell'interpretazione delle emozioni, del verbale, non verbale e così via del nostro interlocutore per darci un'idea di cosa l'altro sta pensando o provando.

In ogni caso, per quanto molto buone e sviluppate, queste abilità non ci danno la certezza di cosa sta realmente pensando l'altro. Certezza che crolla soprattutto quando l'altro riesce a mascherare intenzionalmente ad esempio quello che sta provando.

## Le tecnologie per mappare il cervello: l'EEG

Senza tralasciare tutte quelle situazioni nelle quali abbiamo la necessità di sapere veramente se l'altro sta mentendo oppure no come ad esempio nelle testimonianze.

Ma perché parlare di questo oggi?

Negli anni gli scienziati, neuroscienziati soprattutto, si sono impegnati per mappare il nostro cervello anche per trovare un modo per “leggere” le informazioni in esso contenute.

Questo lavoro certissimo ha permesso di sviluppare delle tecnologie che ci hanno via via avvicinato a esplorare cosa accade nella nostra mente arrivando a tradurre determinati segnali elettrici come specifici stati della mente (ad esempio se siamo concentrati o rilassati).

Questo grazie anche a dispositivi EEG (elettroencefalogramma) che permettono con dei sensori (elettrodi) posizionati sopra la testa su punti specifici di catturare i segnali elettrici del nostro cervello. Tale tecnologia è molto evoluta nel tempo e permette una misurazione dinamica nel tempo dell'attività cerebrale.

## Elon Musk e Neuralink: l'impianto nel cervello

Recentemente sono state sviluppate altre tecnologie come quella che ha realizzato la società Neuralink di Elon Musk. Un progetto che grazie all'uso di un piccolo impianto collocato all'interno del cervello permetterebbe di acquisire dei segnali elettrici e di leggere o intervenire in tempo reale nei vari processi mentali e cognitivi dall'interno.

Stando a quanto dichiarato durante la presentazione avvenuta nel 2022, con lo scopo, di curare problematiche come le lesioni spinali e cerebrali, il recupero della vista come anche problematiche comportamentali o patologie neurologiche. Inoltre, il dispositivo consentirebbe un nuovo tipo di interfaccia cervello computer.

In merito a Neuralink, solo recentemente (Maggio 2023) è stato dato il via libera dalla FDA (Food and Drug Administration) per procedere con la sperimentazione umana di questa tecnologia quindi, ancora, non sappiamo quanto sarà efficace.

Possiamo dire che almeno potenzialmente dovrebbe essere più affidabile e sensibile delle apparecchiature EEG esterne come quelle usate fino ad oggi. Ovviamente il tutto è da prendere con le pinze anche per tutta una serie di questioni etiche e legate alla sicurezza fisica.

## La Risonanza Magnetica Funzionale (fMRI) e l'intelligenza artificiale

Un ulteriore strumento attualmente impiegato per sondare l'attività cerebrale è la Risonanza Magnetica Funzionale (fMRI); un tipo specifico di risonanza magnetica usata sia in ambito clinico che di ricerca per rilevare quali aree del cervello vengono attivate durante alcuni compiti o nel funzionamento normale ma senza l'uso di strumenti invasivi come quelli usati nel progetto Neuralink.

In breve, grazie a questo strumento si possono ricavare delle immagini di aree molto specifiche con un'elevata risoluzione e senza dover usare alcuna sostanza esterna come mezzo di contrasto.

Ma cosa succede se uniamo questa tecnologia con l'Intelligenza Artificiale (IA)?

## L'uso del deep learning per "leggere" i pensieri

In alcune ricerche sono state impiegate tecniche di deep learning per decodificare (interpretare) i segnali catturati dalla fMRI per migliorarne la qualità ma soprattutto “leggere o vedere” all'interno del cervello.

In una ricerca del 2018 [3] gli studiosi grazie all'uso del deep learning potevano comprendere dai segnali cerebrali registrati che cosa i partecipanti alla ricerca stavano vedendo.

I ricercatori ottenevano delle immagini grezze di quello che la persona stava osservando (dei filmati naturali) e anche delle descrizioni testuali. Ad esempio, se la persona stava guardando un video del volo di un uccello, l'immagine ricostruita dal computer era simile a quella di un uccello e la parola suggerita dal modello era appunto “uccello”. L'affidabilità dei modelli impiegati poteva raggiungere il 52% circa delle osservazioni tra i soggetti e fino al 72% circa delle osservazioni intra soggetto.

Forse ancora più impressionante i risultati raggiunti in una ricerca pubblicata a maggio 2023 [4] nella quale gli studiosi hanno indagato se fosse possibile e con quali risultati “leggere” i pensieri dei soggetti coinvolti sempre utilizzando la fMRI e algoritmi di IA.

Come nella ricerca precedente, prima della fase di riconoscimento dei segnali cerebrali, i modelli di IA sono stati allenati tramite sessioni di apprendimento sui segnali generati dai tre partecipanti, tramite la registrazione dell'attività cerebrale dei soggetti sviluppatasi durante l'ascolto di 16 ore di storie.

Dopo la fase di allenamento, i soggetti hanno ascoltato altre storie di test (non usate nella fase precedente) mentre il decodificatore (il modello usato per interpretare i segnali registrati dalla fMRI) traduceva quanto raccolto.

### Le performance del decodificatore

Qui un esempio di una delle frasi test utilizzate “i didn't know whether to scream cry or run away instead i said leave me alone i don't need your help adam disappeared and i cleaned up alone crying”.

Qui invece quello che il decodificatore ha riportato interpretando i segnali cerebrali “started to scream and cry and then she just said i told you to leave me alone you can't hurt me anymore i'm sorry and then he stormed o i thought he had left i started to cry”.

Come spiegato dai ricercatori, nell'esempio sopra come per altri, il decodificatore è riuscito a “leggere” esattamente alcune parole o frasi e comunque a coglierne il senso.

Le performance del decodificatore hanno raggiunto una percentuale tra il 72% e l'82% di affidabilità di quanto ascoltato dai soggetti.

Quindi, grazie alle IA è possibile non solo migliorare la sensibilità di tecnologie come la fMRI ma spingere l'interpretazione dei segnali raccolti quasi a comprendere nel dettaglio quello che la persona sta vedendo, ascoltando o pensando.

Come dichiarato dagli autori grazie a questo studio è stato possibile verificare che questi obiettivi sono potenzialmente raggiungibili grazie a dispositivi non invasivi e soprattutto rispettosi della privacy dei soggetti.

Infatti, come riportato nell'articolo, senza la cooperazione attiva dei soggetti, sia nella fase di addestramento che nella fase di riconoscimento, i risultati del decodificatore crollano drasticamente.

## Conclusioni

In ogni caso è fondamentale aumentare la consapevolezza sia dei benefici dell'uso di queste tecnologie (basti pensare alla possibilità, ad esempio, per quelle persone che non possono comunicare per una lesione cerebrale) ma anche dei rischi di come segnali non correttamente decodificati potrebbero essere usati per scopi malevoli.

## Bibliografia

- [1] Semerari, A., Carcione, A., Dimaggio, G., Falcone, M., Nicolò, G., Procacci, M., & Alleva, G. (2003). How to evaluate metacognitive functioning in psychotherapy? The metacognition assessment scale and its applications. *Clinical Psychology and Psychotherapy*, 10(4).
- [2] Bateman, A. W., & Fonagy, P. (2004). Mentalization-based treatment of BPD. In *Journal of Personality Disorders* (Vol. 18, Issue 1).
- [3] Haiguang Wen and others, Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision, *Cerebral Cortex*, Volume 28, Issue 12, December 2018, Pages 4136–4160, <https://doi.org/10.1093/cercor/bhx268>
- [4] Tang, J., LeBel, A., Jain, S. et al. Semantic reconstruction of continuous language from non-invasive brain recordings. *Nat Neurosci* 26, 858–866 (2023). <https://doi.org/10.1038/s41593-023-01304-9>.

*I quaderni di*

# Agenda **Digitale**

---

## NETWORK **DIGITAL** 360

Network Digital360 è il più grande network in Italia di testate e portali B2b dedicati ai temi della Trasformazione Digitale e dell'Innovazione Imprenditoriale, con oltre 50 fra portali, canali e newsletter.

Ha la missione di diffondere la cultura digitale e imprenditoriale nelle imprese e pubbliche amministrazioni italiane e di fornire a tutti i decisori che devono valutare investimenti tecnologici informazioni aggiornate e approfondite.

Il Network è parte integrante di Digital360HUB, il polo di Demand Generation di Digital360, che mette a disposizione delle tech company un'ampia gamma di servizi di comunicazione, storytelling, pr, content marketing, marketing automation, inbound marketing, lead generation, eventi e webinar.

VIA COPERNICO, 38

20125 - MILANO

TEL. 02 92852785

MAIL: [MARKETING@DIGITAL4.BIZ](mailto:MARKETING@DIGITAL4.BIZ)

©ICT & Strategy

