

I quaderni di
Agenda  **Digitale** ^{eu}

GENNAIO – APRILE 2023

n. 0013

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.

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.

Direzione e comitato editoriale

Direttore responsabile

Alessandro Longo

Direttori scientifici

Paolo Ferri, Mario Morcellini

Comitato editoriale

Giovanni Boccia Artieri, Mario Pireddu, Luca Toschi

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
Lelio Demichelis, Università Insubria, sociologia, economia
Daniela Di Donato, Docente di lettere, Dottoranda di ricerca presso Sapienza Università di Roma- Dipartimento di Psicologia dei processi di sviluppo e socializzazione, Collaboratrice del Crespi
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

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

Affrontare le sfide dell'AI dal lato tecnologico e umano: gli esempi virtuosi	7
Di Marco Aldinucci , Dip. di Informatica, Università di Torino e CINI HPC-KTT, Barbara Cantalupo , Research Engineer Dipartimento di Informatica - Università degli Studi di Torino, Lea Iaia , Senior Researcher (RTDb) of Management at University of Turin · Dipartimento di Informatica - Università degli Studi di Torino, Marco Pironti , Professor of Innovation Management and Entrepreneurship President of ICxT Innovation Center	
Robot coscienti, realtà possibile o utopia? Cosa dicono gli studi	17
Di Antonio Chella , Laboratorio di Robotica, Dipartimento di Ingegneria -Università degli Studi di Palermo	
Giocando s'impara? Alla ricerca delle proprietà educative "elevate" di giochi e videogiochi	25
Di Stefano Triberti , professore associato di Psicologia Generale presso Università Telematica Pegaso e Maria Concetta Carruba , ricercatore in Pedagogia Speciale presso Università Telematica Pegaso	
I limiti e i rischi di ChatGPT non sono quelli che crediamo, possiamo ancora evitare il peggio	33
Di Giuseppe Attardi , professore di Informatica all'Università di Pisa	
Il deepfake può anche aiutare nella lotta al revenge porn	43
Di Marco Viola , Università degli Studi Roma Tre	
Intelligenza artificiale, che ne sarà dell'uomo? Le riflessioni della Chiesa	49
Di Giovanni Tridente , docente di Position Papers alla Pontificia Università della Santa Croce	
La cancel culture non è un'ideologia: proviamo a comprenderla con la semiotica	55
Di Marco Giacomazzi , Università di Bologna e Francesco Barchiesi , Laureato in semiotica	
Un'AI etica è responsabilità di tutti: le basi per realizzarla	66
Di Francesca Rossi , IBM AI Ethics Global Leader, IBM Research, T.J. Watson Research Center, New York, USA	
La demenza digitale esiste? Impariamo a distinguere, con la ricerca "single case"	74
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	
L'interpretazione dei dati: il processo di riappropriazione in quattro fasi	82
Di Maurizio Ferraris , professore ordinario di filosofia teoretica presso la Facoltà di Lettere e Filosofia dell'Università degli Studi di Torino	

Affrontare le sfide dell'AI dal lato tecnologico e umano: gli esempi virtuosi

Le potenzialità delle Intelligenze artificiali dipendono dalla capacità di osservazione, di critica e di analisi di chi è chiamato a svilupparle e a usarle. Occorre una governance in grado di comprensione mirata e di contestualizzazione puntuale, pena un uso limitato e inadeguato delle IA stesse. Su cosa riflettere e perché

Di **Marco Aldinucci**, Dip. di Informatica, Università di Torino e CINI HPC-KTT

Barbara Cantalupo, Research Engineer Dipartimento di Informatica - Università degli Studi di Torino

Lea Iaia, Senior Researcher (RTDb) of Management at University of Turin · Dipartimento di Informatica - Università degli Studi di Torino

Marco Pironti, Professor of Innovation Management and Entrepreneurship President of ICxT Innovation Center

Dalla sua nascita, teorizzata ormai più di sette decenni fa, l'[Intelligenza artificiale](#) ha attraversato momenti di maggiore entusiasmo e altri più complessi, chiamati inverni. **Lo scenario attuale è completamente diverso da quello passato** e, pure non potendo scongiurare il sopraggiungere di un altro inverno, è certamente vero che il tema è diventato più ampio e coinvolge un numero maggiore di tecnologie e attori.

È pertanto un momento cruciale per ripensare l'approccio alle intelligenze artificiali e alle loro possibili evoluzioni, per incoraggiare lo sviluppo tecnologico e infrastrutturale, ma anche per evitare distorsioni sia a livello imprenditoriale - per consentire cioè a tutte le imprese di sfruttarne i vantaggi - che sociale, per far sì che nessuno sia discriminato o resti indietro.

In Italia non mancano esempi virtuosi **su come affrontare le sfide dell'AI sia dal punto di vista tecnologico che del fattore umano.**

Perché ne parliamo proprio oggi?

Le applicazioni dell'IA sono diventate popolari **in ambito industriale e sociale** solo negli ultimi dieci anni, grazie soprattutto a una branca dell'IA [chiamata Machine learning](#). L'insieme delle tecniche di ML hanno dimostrato la capacità di risolvere compiti complessi, come la regressione, la classificazione, il clustering e [filtraggio di dati](#), con prestazioni uguali e talvolta superiori a quelle di operatori umani esperti.

Esistono diversi approcci per l'addestramento dei modelli di ML, fra i quali il cosiddetto [apprendimento supervisionato](#), che consiste nell'addestrare un modello mediante esempi nella forma di possibili input associati ai rispettivi output desiderati, con l'obiettivo di sintetizzare una regola generale che associ un input all'output corretto.

Questo processo è ripetuto per moltissimi esempi, fino alla convergenza del modello ad una configurazione che sia **capace di generalizzare in modo appropriato** gli esempi, cioè di dare risposte plausibili per input che mai visti.

La radice della grande accelerazione di prestazioni osservata per le tecniche di MI è la combinazione di diverse innovazioni che hanno permesso di aumentare la capacità di analizzare e generalizzare grandi insiemi di esempi. Fra queste innovazioni:

- **l'incremento dei dati in forma digitale** disponibili in tutti gli ambiti applicativi, e quindi degli esempi necessari per il procedimento di apprendimento automatico
- **l'ottimizzazione dei metodi** di addestramento dei modelli MI
- l'utilizzo di **acceleratori per il calcolo parallelo** (come le GPU), in grado accelerare in modo eccezionale e a un costo accessibile la velocità nel processare gli esempi e quindi aumentare il numero di esempi che possono essere processati in tempo utile (misurabile in ore, giorni o anche mesi di calcolo).

Come risultato, il MI e più in generale le tecniche di IA hanno permesso di costruire modelli matematici grazie ai quali i calcolatori elettronici riescono ad affrontare in modo accurato e preciso diversi compiti, come **riconoscere oggetti in un'immagine**, comprendere richieste espresse in linguaggio naturale e in diverse lingue, [monitorare anomalie](#) identificabili come rischi, ovvero riconoscere fenomeni individuando delle tendenze.

Quanto descritto porta a prevedere uno sviluppo esponenziale e progressivamente pervasivo dell'IA, con importanti risvolti sul modo di vivere, lavorare, interagire e, addirittura, di **pensare e creare delle persone**, delle imprese, delle istituzioni nonché dei più ampi ecosistemi, economie, rivoluzionando equilibri e meccanismi di creazione di valore nel tempo.

Pmi, il ruolo dell'IA

L'IA può gestire la **complessità dell'impresa**, dalle più semplici funzioni di automazione dei processi all'elaborazione di più complesse previsioni, fornendo una qualità superiore, una maggiore efficienza e risultati migliori nell'elaborazione delle informazioni rispetto alla semplice [elaborazione degli analisti](#). Le informazioni che l'IA mette a disposizione delle imprese, infatti, consentirebbero ai manager di aumentarne la capacità di leggere e interpretare i dati.

Tuttavia, investire e implementare una nuova tecnologia non è mai sufficiente al fine di determinare una migliore performance, poiché **è necessario un ripensamento radicale** di come le imprese possono essere strutturate per sfruttare le opportunità che l'IA offre.

Il successo è una prerogativa delle imprese in grado di integrare tecnologie e capacità manageriali e organizzative che ne supportino la continua trasformazione in nuove idee, ovvero in [prodotti e servizi da portare nel mercato](#). Le criticità collegate a tale aspetto attengono, da un lato, al divario che nasce tra la crescita esponenziale delle prestazioni tecnologiche (quindi **dalla curva di evoluzione tecnologica**) e, dall'altro lato, alla velocità ridotta con cui le innovazioni vengono adottate e assorbite, ovvero fatte proprie da parte di aziende e altre organizzazioni.

Quest'ultima è collegata alla capacità di assorbimento richiesta dall'IA, associata al raggiungimento di una comprensione di base che permetta al management di utilizzarle proficuamente.

Nonostante il potenziale dell'IA, le Pmi che caratterizzano la conformazione imprenditoriale italiana spesso **non posseggono le risorse** necessarie per poterla implementare nei propri processi produttivi. La causa può risiedere nell'assenza di capitale e, più sovente, nella impossibilità di trovare le competenze idonee per comprendere il ruolo che l'IA può avere nell'impresa, nella dinamica uomo-macchina.

Dall'Intelligenza artificiale all'Intelligenza aumentata

Se l'Intelligenza artificiale sarà realmente un dirompente fenomeno economico, sociale, ambientale e non una mera tecnologia, **dipenderà dai punti di osservazione e di interpretazione** perché tante dovranno essere le prospettive di analisi, anche critiche, che bisognerà adottare per leggere correttamente e sistemicamente cosa stiamo vivendo e soprattutto cosa ci attende.

Bisognerà declinare le **Intelligenze artificiali** al plurale, poiché partire dalla comprensione dei diversi obiettivi e delle tecniche computazionali cui fanno riferimento [ne consente una valorizzazione più olistica](#) e human-centered.

In tale prospettiva, non ci si concentra esclusivamente sulla **contrapposizione del carattere artificiale al naturale** (o ancora meglio umano), ma si enfatizza la capacità delle IA di posizionare l'uomo al centro del proprio divenire, anche tecnologico, e aumentarne talune capacità e potenzialità.

Per chi scrive, il compito delle IA è la complementarietà tra tecnologia ed essere umano e questa, a sua volta, corrobora il legame tra i due attori conducendo così a un'ulteriore evoluzione del concetto stesso di IA, ovvero l'intelligenza aumentata. Quest'ultima è un costrutto ancora in fase di definizione nella letteratura accademica per via della sua complessità ed evoluzione, [è considerata lo step successivo](#) nello sviluppo della ricerca sull'IA.

L'intelligenza aumentata

L'intelligenza aumentata si ricollega a un approccio progettuale e un'implementazione dell'IA che non sostituisce ma migliora l'intelligenza umana. L'uomo fa parte del sistema di IA e grazie a ciò si crea un rapporto che espande le capacità umane di elaborazione delle informazioni in una dimensione che si colloca nel mezzo tra la realtà e la virtualità. Secondo Matthew Sadiku del Roy G. Perry College of Engineering (Texas) l'intelligenza aumentata [segue un ciclo di cinque fasi](#), così articolato:

- **Understanding:** i dati vengono utilizzati per accrescere i sistemi che li esaminano ed estrapolano significato
- **Interpretation:** nuovi dati vengono confrontati dal sistema con dati preesistenti in modo da interpretarli
- **Reasoning:** il sistema crea "output" o "risultati" per il nuovo set di dati
- **Learn:** l'output al punto precedente viene esaminato dagli esseri umani, che danno un feedback al sistema, consentendogli di adattarsi
- **Assure:** la sicurezza e la conformità delle soluzioni elaborate sono garantite utilizzando la tecnologia blockchain o IA.

Le sinergie derivanti dalla collaborazione tra macchine e uomo richiedono il **supporto proveniente invero da ulteriori attività** che riguardano la selezione di talenti con un profilo di competenze linguistiche e tecniche di dominio, la definizione di una governance adeguata sull'automazione e

sulle persone, la definizione di un management che assicuri una adozione agevole delle nuove modalità di lavoro e potenziamento delle competenze.

In ultimo, non per importanza, le organizzazioni che intendono implementare l'intelligenza aumentata dovrebbero considerare due aspetti rilevanti:

- che tecnologia e processi abbiano un input umano continuo
- che esistano degli **algoritmi white-box** che spieghino il mal funzionamento di alcuni processi.

Al riconoscimento di schemi e all'adattamento di curve (ovvero la ricerca di una relazione che spieghi un insieme esistente di punti dati) offerta dall'IA si contrappone la capacità unica dell'uomo di essere creativo e innovativo. Le IA non sono in grado di assolvere tale compito **poiché non potranno mai sostituirsi completamente all'uomo** a causa dei conflitti etici legati ai possibili bias che verrebbero a crearsi.

Una semplice e significativa equazione proposta da Sadiku identifica immediatamente il concetto di intelligenza aumentata come: *Augmented Intelligence = Human + Computer*

Eppure, secondo chi scrive, tale lettura **non considera l'evoluzione delle capacità**, delle competenze, delle conoscenze, delle tecniche e delle tecnologie che contribuiscono alla trasformazione delle IA così come all'aumento delle potenzialità dell'uomo in un circolo virtuoso e sostenibile, che qui si amplia proponendo una nuova chiave di interpretazione. Infatti, il valore globale dell'intelligenza aumentata include sia la capacità umana di utilizzare in modo avanzato le aumentate capacità fornite dalle IA sia la capacità dell'AI di continuare a consolidare in modo artificiale le nuove conoscenze acquisite, utilizzando proprio il combinato di capacità artificiali e umane.

Attivare un circolo virtuoso per l'IA

Considerato che le IA consentirebbero alle organizzazioni di conseguire prestazioni elevate e fino a ora sconosciute, tanto da essere considerate le tecnologie principali a livello strategico, è altresì vero che le Pmi che compongono il nostro tessuto imprenditoriale nazionale **potrebbero ricavare innumerevoli benefici** dall'implementazione di tali tecnologie soltanto qualora riuscissero a superarne le criticità.

Difatti, le Pmi non sempre dispongono delle risorse - non soltanto economiche e finanziarie, ma soprattutto umane - per poter accedere e adottare tale tecnologia. Allo stesso modo, la ricerca applicata fa un gran lavoro per poter supportare lo sviluppo e l'innovazione delle IA, ma tale ruolo non può essere assolto in autonomia.

Oggetto di dibattito, dunque, diventa non l'ambito strettamente tecnologico dell'IA, **ma quale governance e soprattutto management** si rendono necessari per compensare il rischio che le tecnologie emergenti non tengano conto delle reali capacità dell'uomo di adottarle in maniera consapevole, dell'etica e del valore (non solo economico) in grado di garantire un circolo virtuoso per uno sviluppo eco-sistemicamente sostenibile nel lungo periodo.

Perché governare le IA

In generale, la necessità di una governance nasce in contesti complessi caratterizzati dalla presenza di molteplici attori che entrano in relazione tra loro con diversi obiettivi, risorse e tempistiche, influenzando il raggiungimento di un equilibrio dinamico. L'Intelligenza artificiale impatta su tutte queste componenti, poiché:

- rivoluziona le dinamiche degli attori coinvolti, creando nuovi soggetti, decretando il fallimento o la non sostenibilità e sopravvivenza di altri, cambiando ruoli e strumenti di persone, imprese e istituzioni
- modifica le modalità relazionali e di interazione tra gli attori, ne stravolge il timing, le risorse necessarie, le competenze richieste e i media utilizzati
- genera nuovi equilibri, non più limitati a obiettivi di natura economica e finanziaria ma ampliandosi ad una sostenibilità ambientale, sociale e circolare.

In una prospettiva che contempra l'intelligenza aumentata, l'esigenza di governare tale rivoluzione è ampiamente visibile. Se l'obiettivo delle IA è risolvere problemi e gestire le attività dapprima assegnate agli esseri umani perseguendo i principi di velocità, razionalità, efficienza e scalabilità, l'intelligenza aumentata si pone obiettivi ben più complessi. Quest'ultima, infatti, **mira a integrare sinergicamente idee, creatività, progettualità e azioni umane** amplificando, come già espresso, le potenzialità dell'uomo, che grazie ad essa può avvalersi di risorse che nella quantità e qualità (si pensi ai dati) non sarebbe in grado di gestire razionalmente e in autonomia.

La complessità dello scenario continua ad aumentare focalizzandosi sull'evoluzione degli attori coinvolti (con processi sempre più evidenti e rilevanti di convergenza tra settori economici), la **contaminazione dei saperi e delle competenze**, nonché la ridefinizione dei ruoli. Si pensi, per esempio, al cliente che diventa parte del processo di design, di produzione, di distribuzione o di riuso o al cittadino che diventa parte attiva e partecipativa delle scelte di governo locale, centrale o sovranazionale.

Il focus della governance **non è la tecnologia, ma la capacità dell'uomo di intercettare**, metabolizzare e usare tutte le potenzialità che le nuove tecnologie emergenti sono in grado di offrire.

Perché gestire le IA

I classici momenti manageriali (pianificazione, organizzazione, esecuzione, direzione e controllo) **stanno mostrando tutti i loro limiti** nella loro accezione più tradizionale legata a mercati stabili, continui, con obiettivi definiti e sistematici, risorse quantificate e determinate.

La [trasformazione digitale](#) sta lasciando sempre minori punti di riferimento certi e predefiniti, determinando spesso strategie di impresa non chiare, non correttamente comunicate all'interno e all'esterno delle organizzazioni e non supportate da risorse disponibili adeguate.

La funzione manageriale diventa sempre più magmatica, multidisciplinare, eclettica e dinamica se si considera che:

- la **selezione delle risorse umane** verte sempre più sulle capacità di apprendimento e adattamento piuttosto che sulle competenze delle persone, imponendo un ripensamento di modelli di formazione non solo ex ante ma anche e soprattutto in itinere

- le organizzazioni tendono sempre più a modelli fluidi, che richiedono nuovi modelli di **delega e di valutazione per obiettivi**
- i contesti organizzativi diverranno sempre più “aperti”, in cui il confine tra chi e cosa è “dentro le mura” e cosa è “fuori le mura” è labile, fluido e dinamico. **Le nuove leve di vantaggio competitivo saranno determinate** dalle capacità non più di proteggersi, ma da quelle di condivisione, collaborazione e cooperazione
- la cultura dell’efficienza legata all’ottimizzazione di processi standardizzati lascerà il campo alla cultura dell’innovazione. Ciò sposterà progressivamente l’attenzione dalla mera soluzione del problema **alla capacità di proporre nuove sfide**
- nella pratica manageriale soltanto ciò che è misurabile esiste ed è migliorabile. Tuttavia, in azienda esistono diversi elementi e fattori non visibili. Il contributo delle IA in tali ambiti può essere rappresentato dalla identificazione e formalizzazione di fenomeni non visibili da un occhio umano e razionale (per esempio, la **predictive maintenance** permette di intervenire in maniera programmata, prevenendo un disservizio) ma soprattutto di fornire strumenti e tecniche per misurare in maniera attendibile, veloce e affidabile ciò che in passato non era possibile identificare (per esempio, la [business intelligence](#) basata su data analytics).

Dal punto di vista organizzativo e funzionale non sarà sufficiente individuare un cosiddetto Chief Artificial Intelligence Manager, ma comprendere principi e criteri che definiscano le capacità, le competenze e i requisiti più ampi che deve possedere una vera e propria organizzazione estesa in grado di aumentare le proprie potenzialità, efficacia ed efficienza, **utilizzando le intelligenze artificiali e rimettendo le persone al centro**.

Il principale vantaggio competitivo risiederà nel come le nuove tecnologie consentiranno di sviluppare l’intelligenza emotiva. La nuova era dell’**intelligenza artificiale emotiva** dovrà, infatti, ricostruire contesti non artificiosamente intelligenti, ma che tengano conto di emozioni, stati d’animo, empatia. In altre parole, un’umanizzazione delle tecnologie.

L’impatto del management sulle IA assume ben più ampia valenza se si considera necessariamente la sua condizione di sostenibilità, dai modelli di business che devono adeguarsi ai nuovi modelli che l’evoluzione tecnologica richiederà. Anche se la valutazione di sostenibilità si applica alle imprese inserite in un processo di circolarità, l’adozione sempre più pervasiva delle IA nelle organizzazioni **affretterà la compatibilità e sinergia con le regole e i principi dell’economia circolare**. Oggi aspetto distintivo, domani (non troppo lontano) condizione di sopravvivenza nel mercato.

Non si tratterà solo di management d’impresa, ma di un processo co-evolutivo di management dei sistemi complessi per l’impatto sugli ecosistemi, sia come evoluzione di quelli esistenti e sia di creazione di quelli emergenti.

Un approccio inclusivo e sistemico alle IA

Quanto finora descritto dimostra che le IA vanno affrontate come un complesso e complesso tema di **governance e soprattutto di management**, con un approccio inclusivo e sistemico.

Comprendere tale passaggio, interrogandosi sulle infrastrutture tecnologiche e umane che possano indirizzare e guidare l’evoluzione delle intelligenze artificiali e aumentate, è di fondamentale importanza, poiché **le IA potranno diventare una reale innovazione** se in esse non si cercherà un carattere di resilienza ma di evoluzione sostenibile, rigenerativa, collaborativa e diffusa.

Le IA **non dovranno riportare i sistemi ai vecchi equilibri**, ma dare la forza all'uomo di definire nuovi obiettivi e condizioni di equilibrio determinando innovativi processi eco-evolutivi per le persone, per le imprese, le istituzioni e, a macro-livello, per gli ecosistemi, le economie, gli equilibri e i meccanismi di creazione di valore nel tempo.

Tra gli esempi notevoli di infrastrutture tecnologiche abilitanti per la ricerca in IA sul territorio italiano si distinguono [HPC4AI](#) (High-Performance Computing for Artificial Intelligence) e [ICSC](#) (Centro Nazionale di Ricerca in HPC, Big Data e Quantum Computing).

Esempi di diversa scala e con diversi obiettivi, ma con un comune tratto di innovazione nell'ambito del supporto e della **gestione dei progetti IA in molti e diversi ambiti**: capire il DNA, creare farmaci personalizzati, studiare l'evoluzione del clima, sintetizzare materiali meno inquinanti, sfruttare meglio le fonti energetiche.

HPC4AI: Un living lab per le IA

Un esempio di come affrontare le sfide dell'AI sia dal punto di vista tecnologico che del fattore *umano* è rappresentato dal progetto HPC4AI nato nel 2018 da un'iniziativa congiunta tra l'Università di Torino e il Politecnico di Torino. Il progetto, finanziato con 4,5 milioni di euro grazie al bando competitivo grandi infrastrutture EU POR-FESR 2014-2020 e con il cofinanziamento dei due Atenei, è nato con l'obiettivo di **realizzare una federazione di infrastrutture di calcolo ad alte prestazioni** per scopi di ricerca finalizzate ad applicazioni per l'analisi dei dati, sia nell'ambito dell'intelligenza artificiale che dei Big data.

La componente dell'Università di Torino, HPC4AI@UniTO, ha assorbito i due terzi del finanziamento e ha portato alla **realizzazione di un data center funzionale** allo sviluppo di nuove applicazioni di intelligenza artificiale, in grado di anticipare i trend su questo fronte in un contesto di laboratorio e con un'infrastruttura all'avanguardia.

HPC4AI@UniTO si propone come **un living lab (Open access) per docenti e studenti**, mettendo a disposizione dei ricercatori di diversi domini scientifici una ingente potenza di calcolo al servizio dei metodi delle IA, che possa essere utilizzata facilmente mediante gli strumenti del cloud computing. Al tempo stesso si propone di collaborare con le realtà imprenditoriali del territorio per aumentare la loro capacità di innovare su tecnologie e applicazioni basate sui dati, offrendo accesso a risorse di calcolo e storage in tutto lo spettro dei modelli di servizio cloud, inclusi i servizi sperimentali.

Dal punto di vista infrastrutturale, la prima versione di HPC4AI@UniTO è stata completata a fine 2019, ma grazie al fatto che l'infrastruttura si autofinanzia, oggi è quasi raddoppiata in dimensione e conta oltre 11.000 core con oltre 120 GPU specializzate per workload di IA.

Il laboratorio ruota intorno ad un **green data centre da 250kW**, con un'efficienza energetica allo stato dell'arte (circa il 95% contro il 65-70% della media). La progettazione ha richiesto mesi di lavoro e lo sviluppo di nuove soluzioni per il contenimento del consumo energetico e del rumore, ma il risultato è che il data center costituisce il vero cuore pulsante del laboratorio, posizionato in una teca di vetro al centro del dipartimento di informatica dell'Università di Torino visibile a tutti a dimostrazione che il cloud non è in un posto indefinito, ma nel mezzo della azione della ricerca.

La piattaforma di calcolo si compone di [sistemi modulari cloud-HPC](#), in cui ciascuno modulo è ottimizzato secondo le proprie specificità e poi integrato con la controparte. In particolare, si tratta

di due isole cloud-HPC, uno di produzione e di sviluppo, in modo da poter mettere a disposizione degli utenti un'isola di risorse ad alta affidabilità e sicurezza, **attualmente con zero incidenti** "system-down" in tre anni, e un'isola completamente a disposizione dei ricercatori per sviluppare e sperimentare nuove soluzioni, strumenti e componenti hardware.

Al momento su HPC4AI@iTO sono operativi **oltre 60 progetti di ricerca accademica e industriale**, di cui almeno 10 finanziati da bandi [Horizon 2020](#), con una media di 180 macchine virtuali attive. Molti progetti sono di ambito medico e usano [tecniche di Deep learning](#) con centinaia di Terabyte di dati provenienti da trial delle cliniche universitarie di Torino.

Grazie a un finanziamento competitivo del ministero dello Sviluppo economico di 14,5 milioni di euro, presto verrà aperto un nuovo laboratorio, **HPC4AI-Compute Continuum**, in cui la ricerca sarà focalizzata su [federated learning](#) e analisi del dato, ovvero dove viene prodotto, concentrandoci sulla riservatezza, una caratteristica fondamentale dei dataset di valore (medici, strategici, eccetera).

ICSC: il Centro Nazionale HPC, Big Data e Quantum Computing

Un esempio virtuoso di come rispondere alle necessità infrastrutturali emergenti delle applicazioni IA è rappresentata dal **Centro Nazionale HPC, BigData e Quantum Computing**, [nato nell'estate 2022](#) assieme ai cinque Centri nazionali per la ricerca in filiera previsti dalla Componente "dalla ricerca al business" della Missione "Istruzione e Ricerca" del Piano nazionale di ripresa e resilienza (Pnrr) grazie a 1,6 miliardi di euro.

Tra questi ICSC, il Centro Nazionale di Ricerca in High-Performance Computing, Big Data e Quantum Computing conta 51 membri fondatori distribuiti su tutto il territorio nazionale, provenienti dai settori pubblico e privato, dal mondo della ricerca scientifica e dell'industria.

I Centri nazionali sono **aggregazioni di università, di enti e organismi pubblici e privati** di ricerca, di imprese presenti e distribuite sull'intero territorio nazionale e sono organizzati con una struttura di governance di tipo Hub&Spoke, con l'Hub deputato a svolgere attività di gestione e coordinamento mentre gli Spoke si occupano della ricerca.

ICSC conterà su un finanziamento proveniente da fondi Next Generation EU nell'ambito della Missione Istruzione e Ricerca del Pnrr coordinata dal ministero dell'Università e della ricerca (MUR), pari a **320 milioni di euro, di cui il 41% sarà investito al Sud**. In particolare, del finanziamento complessivo, oltre 100 milioni di euro saranno dedicati al personale, un investimento che viene considerato prioritario, con una partecipazione femminile di almeno il 40% e con quasi 16 milioni di euro riservati all'alta formazione e alla carriera dei giovani.

ICSC aggrega le comunità scientifiche italiane di eccellenza in dieci diversi ambiti, sarà strutturato su due colonne portanti di eguale rilevanza, le infrastrutture e le aree tematiche, e sarà organizzato secondo il modello Hub&Spoke. Il Centro prevede anche il **coinvolgimento delle aziende italiane** per costruire una sinergia tra le comunità scientifiche e il mondo industriale, a beneficio sia del sistema della ricerca sia del sistema produttivo.

L'Hub avrà la responsabilità di validare e gestire i programmi di ricerca, le cui attività verranno elaborate e realizzate dagli Spoke e dalle realtà a essi affiliate, anche attraverso bandi aperti a istituzioni di ricerca e aziende. Gli Spoke saranno dieci, due di questi avranno una natura fortemente fondazionale (Future HPC e Quantum Computing), mentre otto saranno concentrati sulle applicazioni HPC.

Specificamente, lo Spoke FutureHPC aprirà due laboratori nazionali sulle tecnologie hardware (presso Università di Bologna) e **software** (presso l'Università degli Studi di Torino). Quest'ultimo, chiamato "Software & Integration lab" (**SWI**), sorgerà presso il Dipartimento di Informatica dell'Università di Torino e servirà da contamination lab, cioè **da spazio di co-working dove le quindici Università** dello Spoke FutureHPC lavoreranno insieme alle 8 grandi aziende per sviluppare le tecnologie abilitanti del futuro e per formare esperti in ambito cloud-HPC e AI-HPC con una chiara visione dello stato dell'arte della ricerca nel settore e delle esigenze in ambito industriale. Fra le aziende coinvolte nel laboratorio SWI: Eni, Intesa SanPaolo, Leonardo Company, Unipol, Thales Alenia, Autostrade, Fincantieri e Sogei.

Il Centro **svolgerà attività di Ricerca e Sviluppo**, a livello nazionale e internazionale, a favore dell'innovazione nel campo delle simulazioni, del calcolo e dell'analisi dei dati ad alte prestazioni. Queste attività saranno svolte a partire da una infrastruttura d'avanguardia a livello internazionale per l'High Performance Computing e la gestione dei Big data, capace di mettere a sistema le risorse e di promuovere e integrare le tecnologie emergenti.

Nei prossimi anni ICSC implementerà soluzioni che **porteranno a una velocità di rete superiore a 1 Terabit/secondo**, e metterà a disposizione degli utenti una infrastruttura cloud tale da consentire la gestione di attività alla frontiera nella ricerca scientifica e nello sviluppo industriale. Il nuovo Centro Nazionale di supercalcolo si focalizzerà da una parte sul mantenimento e il potenziamento dell'infrastruttura HPC e Big data italiana, e dall'altra sullo sviluppo di metodi e applicazioni numeriche avanzati e di strumenti software, per integrare il calcolo, la simulazione, la raccolta e l'analisi di dati di interesse per il **sistema della ricerca e per il sistema produttivo e sociale**, anche attraverso approcci cloud e distribuiti.

ICSC coinvolgerà e promuoverà le migliori competenze interdisciplinari delle scienze e dell'ingegneria, permettendo innovazioni sostanziali e sostenibili in campi che vanno dalla ricerca di base alle scienze computazionali e sperimentali per il clima, l'ambiente, lo spazio, dallo studio della materia e della vita alla medicina, dalle tecnologie dei materiali ai sistemi e ai dispositivi per l'informazione. Il Centro sosterrà l'alta formazione e promuoverà lo sviluppo di politiche per la gestione responsabile dei dati in prospettiva di open data e open science, **coniugando profili di regolamentazione, standardizzazione e compliance**. ICSC sarà un'infrastruttura cloud/HPC condivisa e aperta, e rappresenterà un asset strategico unico per l'Italia, ma anche per la comunità internazionale.

Il Centro Nazionale ICSC è entrato in piena attività dal 1° settembre 2022. Gli Spoke e i loro affiliati saranno impegnati a regime, ossia a partire dal 2026, a sostenere sul proprio budget ordinario alcune misure finanziate in fase di start-up dai fondi Pnrr 2022-2025 e a finanziare il Centro con un contributo annuale a partire dal 2022.

Il Centro intende assolvere al suo ruolo strategico **moltiplicando le opportunità per partner e stakeholder**, massimizzando l'impatto socioeconomico in un contesto competitivo, implementando qualcosa che oggi non esiste, senza sovrapporsi o competere con i singoli istituti di ricerca già presenti, ma creando invece valore aggiunto per i propri partner e per la società. E la chiave del successo dell'iniziativa risiede nella capacità di creare queste condizioni già nella fase di start-up.

La nostra è la società dei dati. Negli anni più recenti, la sempre maggiore digitalizzazione di pressoché ogni attività umana ha prodotto e messo a disposizione un'immensa quantità di dati e di informazioni e, nel futuro prossimo, questa tendenza andrà sempre più incrementandosi e producendo, già nei prossimi anni, moli di dati senza precedenti. Archiviare, ordinare, condividere, processare e interpretare questi dati, i cosiddetti Big data, è diventata **la grande impresa di oggi**

per rendere competitivo il Paese, sia nella ricerca scientifica sia nel sistema produttivo, così da essere in grado di affrontare le sfide presenti e future su questioni chiave per la nostra società e il nostro pianeta, come la salvaguardia dell'ambiente, la tutela della salute personale e pubblica, la realizzazione di smart city, lo sviluppo sostenibile.

La gestione dei Big data è quindi una priorità, per la quale sono necessarie ingenti risorse di calcolo, lo sviluppo di soluzioni tecnologiche innovative, come i computer quantistici, e di algoritmi e modelli matematici e statistici sempre più avanzati. In questo contesto, il supercalcolo, le simulazioni numeriche, l'intelligenza artificiale, il machine learning e le tecnologie di frontiera sono **strumenti essenziali per incentivare sviluppi e scoperte** sia del mondo industriale sia del mondo della ricerca scientifica e, di conseguenza, per la crescita economica e culturale della nostra società.

ICSC ambisce a diventare un ecosistema attrattivo a livello internazionale, rafforzando la competitività italiana e **contribuendo alla leadership dell'Europa nel settore**. Questo è lo scenario nel quale nasce il Centro Nazionale HPC, Big Data e Quantum Computing, con il triplice scopo di costruire un'infrastruttura di supercalcolo italiana, di aggregare le risorse di ricerca e di innovazione nei settori maggiormente strategici per il Paese, e di posizionarsi come la piattaforma nazionale a supporto di iniziative scientifiche e industriali.

Bibliografia

A. Agrawal, J. Gans, A. Goldfarb, «Exploring the impact of artificial intelligence: prediction versus judgment», n. 24626, NBER Working Paper, NBER Working Paper, Cambridge, MA, 2018.

J. Bughin, E. Hazan, S. Lund, P. Dahlström, A. Wiesinger, A. Subramaniam, «Skill shift: automation and the future of the workforce», McKinsey Global Institute, 2018.

M. Iansiti, K.R. Lakhani, *Competing in the age of AI: strategy and leadership when algorithms and networks run the world*, Cambridge, Harvard Business Press, 2020.

M.G. Jacobides, S. Brusoni, F. Candelon, «The Evolutionary Dynamics of the Artificial Intelligence Ecosystem», *Strategy Science*, 6(4), 2021, pp. 412-435.

M. Wójcik, «Augmented intelligence technology. The ethical and practical problems of its implementation in libraries», *Library Hi Tech*, 39(2), 2021, pp. 435-447.

M.N.O. Sadiku, T.J. Ashaolu, A.A. Majebi, S.M. Musa, «Augmented Intelligence», *International Journal of Scientific Advances (IJSCIA)*, 2(5), Sep-Oct, 2021, pp. 772-776.

Pironti M. (a cura di), *Intelligenze artificiali e aumentate*, Milano, Egea, 2022.

M. Aldinucci, S. Rabellino, M. Pironti et al., HPC4AI, an AI-on-demand federated platform endeavour, *ACM Computing Frontiers*, Ischia, Italy, 2018. doi:10.1145/3203217.3205340.

UNITO news, «PNRR: nasce il Centro Nazionale di Supercalcolo», luglio 2022, <https://www.unito.it>.

Robot coscienti, realtà possibile o utopia?

Cosa dicono gli studi

C'è un fervente dibattito scientifico attorno alla coscienza, continuamente ravvivato da chi è convinto di poterla emulare e da chi, al contrario, si dice certo che sia funzione unicamente biologica. Le ricadute degli studi che tendono a stabilirlo impattano anche sulla definizione del concetto di persona

Di **Antonio Chella**, Laboratorio di Robotica, Dipartimento di Ingegneria -
Università degli Studi di Palermo

antonio.chella@unipa.it

I recenti successi dei sistemi conversazionali come **ChatGPT** e dei sistemi di generazione automatica di immagini a partire da descrizioni testuali come DALL-E hanno riaperto il dibattito sulla possibilità di costruire robot coscienti; si veda ad esempio il recente articolo apparso sul New York Times [1].

Lo studio della bioingegneria e della bioinformatica sta dando un grande aiuto alla comprensione dei processi biologici e delle neuroscienze che sono alla base della coscienza. Oggi abbiamo accesso a tanti dati del cervello, provenienti dalla risonanza magnetica, dalla TAC, dall'elettroencefalogramma. La grande iniziativa europea legata allo **Human Brain Project**, sebbene criticata per diverse ragioni, e che aveva originariamente lo scopo di replicare in maniera minuziosa il cervello umano mediante una modellazione dei singoli neuroni, ha innegabilmente dato un grande impulso a questi studi.

Il campo multidisciplinare della ricerca sulla coscienza artificiale mira a indagare il fenomeno della coscienza alla luce della robotica e dei sistemi artificiali, della psicologia, della filosofia della mente, dell'etica e delle neuroscienze. Il presente contributo non vuole e non può essere una rassegna esaustiva del **ricco dibattito filosofico** riguardante la possibilità di costruire effettivamente robot coscienti, ma una discussione su alcuni modelli computazionali che traggono ispirazione dagli studi sulla coscienza.

Finora non esiste una definizione accettata di coscienza: per una ampia rassegna sulle definizioni finora proposte si veda Vimal [2]. Secondo Searle [3] la coscienza consiste in stati e processi interiori, qualitativi e soggettivi, di consapevolezza. La coscienza, sempre secondo Searle, inizia quando ci svegliamo al mattino da un sonno senza sogni e continua fino a quando ci addormentiamo di nuovo, moriamo o entriamo in coma.

Sebbene ci siano **posizioni filosofiche contrastanti**, è utile sottolineare la distinzione della coscienza intesa come *esperienza* rispetto alla coscienza come *funzione* (si veda Block [4]). Sulla base della coscienza intesa come esperienza, un soggetto è cosciente quando prova esperienze visive, sensazioni corporee, immagini mentali, emozioni. Secondo Chalmers [5], il problema difficile della coscienza è proprio quello di spiegare queste esperienze interne. Secondo Nagel [6], risolvere il problema difficile della coscienza equivale a spiegare “cosa si prova ad essere un pipistrello.”

Secondo la coscienza intesa come funzione, un soggetto cosciente, tra le altre cose, è capace di integrare informazioni ed elaborarle globalmente [7], è introspektivamente consapevole di sé [8], possiede un modello di sé stesso e dell'ambiente esterno [9], può anticipare attività percettive e comportamentali [10], agisce interagendo con il mondo esterno [11].

Cenni storici

La prima occorrenza della parola *coscienza artificiale* si trova nel libro ***Cybernetic Machines*** di **T.N. Nemes**, pubblicato in Ungheria nel 1962. Il libro fu tradotto in inglese nel 1970. L'autore considera la coscienza artificiale come la capacità di un robot di discriminare sé e gli altri. Nemes propone un circuito capace di distinguere tra segnali interni che generano frasi come *Io vado*, da circuiti di riconoscimento della forma e percezione del movimento che elaborano segnali provenienti dall'esterno e capaci di generare frasi come *Egli va*.

Il moderno quadro scientifico della coscienza artificiale è stato introdotto da Igor Aleksander che, sin dal 1992, ha proposto un quadro di riferimento scientifico per gli studi sulla coscienza nei robot [12].

Il simposio su ***Can a Machine Be Conscious***, organizzato dalla Swartz Foundation nel 2001, è stata un'altra pietra miliare per la coscienza dei robot. Le osservazioni conclusive di Christof Koch, valide ancora oggi, affermavano che *non conosciamo nessuna legge o principio fondamentale operante in questo universo che proibisca l'esistenza di sentimenti soggettivi in artefatti progettati o evoluti dall'uomo* [13].

Dal 2001 ad oggi, molte conferenze, workshop, numeri speciali di riviste sono stati dedicati al campo della coscienza dei robot. In particolare, durante l'estate del 2017, il prestigioso Stanford Research Institute in California ha organizzato una serie di workshop su *Technology and Consciousness* i cui atti forniscono un buon riassunto dello stato dell'arte della ricerca nel settore [14].

Una fonte continua di informazioni in questo filone di ricerca è il *Journal of Artificial Intelligence and Consciousness* <https://www.worldscientific.com/worldscinet/jaic>

Neuroscienze e robotica

La coscienza è un importante argomento di ricerca nelle neuroscienze. Le neuroscienze contribuiscono al dibattito sui robot coscienti in quanto i modelli biologici ispirano i modelli computazionali alla base della coscienza. Lo scomparso premio Nobel Gerald Edelman è stato un precursore nell'uso dei robot per analizzare le teorie sull'organizzazione del cervello e sulla coscienza biologica.

Edelman e colleghi hanno creato la serie di robot **Darwin** [15] per simulare alcuni aspetti della coscienza. I robot incorporano modelli di modifiche sinaptiche per generare il comportamento secondo il contesto e la storia e senza la necessità di programmare il robot stesso. Tra questi, Darwin III è un modello di robot che lavora in un ambiente simulato ed è in grado di imparare la coordinazione sensorimotoria, la capacità di seguire gli oggetti, la capacità di raggiungere e afferrare gli oggetti e di categorizzarli interagendo con l'ambiente. **Darwin VII** [16] può effettuare la categorizzazione percettiva e le risposte condizionate in semplici compiti di ricerca del cibo; Darwin VIII può analizzare una scena percepita e formare categorie adeguate, senza la necessità di un sistema di controllo.

Il neuroscienziato **Giulio Tononi** ha proposto la *Teoria dell'Informazione Integrata* della coscienza (si veda Oizumi et al. [17] per una versione aggiornata della teoria). Questa teoria è oggi la più dibattuta e analizzata, e ne derivano importanti risultati anche per i robot. La formulazione originale parte dall'osservazione che l'esperienza cosciente è differenziata perché il repertorio potenziale dei diversi stati coscienti è enorme. Allo stesso tempo, l'esperienza cosciente è integrata, poiché ogni stato cosciente è percepito come una singola entità. Quindi, il substrato dell'esperienza cosciente deve essere un'entità integrata in grado di differenziare tra un enorme repertorio di stati diversi.

Secondo questa teoria, l'esperienza è una grandezza fondamentale della natura come la massa, la carica, l'energia. Qualsiasi sistema fisico può avere esperienze nella misura in cui è capace di integrare informazioni. Tuttavia, come discusso da Koch [18], la teoria suggerisce che i computer convenzionali non sono in grado di eseguire un'efficace integrazione delle informazioni, e non possono essere quindi in grado di avere esperienze. Un robot può soltanto essere un'entità che imita il comportamento cosciente ma incapace di provare effettivamente qualcosa.

Modelli computazionali

Un percorso di indagine nella coscienza dei robot consiste nell'individuare un insieme minimo di caratteristiche che devono essere verificate per affermare che un robot sia effettivamente cosciente. Aleksander e Dunmall [19] hanno proposto cinque assiomi che dovrebbero essere verificati da un organismo cosciente. Essi sono: (i) un organismo che non apprende non può essere cosciente; (ii) un organismo cosciente possiede uno stato interno in grado di rappresentare il mondo esterno; (iii) un organismo cosciente è in grado di prestare attenzione ai contenuti del suo stato interno; (iv) un organismo cosciente è in grado di generare stati interni relativi a sequenze di segnali esterni e di generare azioni adeguate; (v) l'organismo è in grado di prevedere eventi esterni attraverso sviluppi controllati del suo stato interno. Aleksander e Dunmall hanno tradotto questi assiomi in termini di vincoli matematici che devono essere soddisfatti da un sistema neurale per essere considerato come dotato di una coscienza minima. Aleksander [20] ha successivamente proposto un modello computazionale derivato dagli assiomi e implementabile in un robot.

La teoria dello spazio di **lavoro globale** (Global Workspace Theory - GWT) è stata proposta da Baars [21] come unificazione di diversi processi nella corteccia cerebrale. Baars ha osservato che il cervello può eseguire un'enorme quantità di elaborazioni parallele inconscie, mentre la coscienza è seriale ed ha capacità limitata. La GWT si basa sull'ipotesi che il cervello sia un insieme di processori specializzati che lavorano in maniera inconsapevole. **La coscienza è associata a uno spazio di lavoro globale** il cui contenuto viene trasmesso a tutti i processori. I processori lavorano in parallelo e competono per accedere allo spazio di lavoro globale. Quando un processore vince questa competizione e ottiene l'accesso allo spazio di lavoro globale, entra nella coscienza, recluta gli altri processori, e seleziona l'azione corrispondente.

Consideriamo, per esempio, **un agente che assiste a una scena con oggetti in movimento**. Secondo la GWT, ogni oggetto in movimento è elaborato da un processore, e tutti i processori competono per ottenere l'accesso allo spazio di lavoro globale. A un certo punto, un processore, corrispondente ad esempio a una palla che si muove verso l'agente, vince la competizione. La palla che si muove entra allora nella coscienza e il relativo processore recluta altri processori per selezionare l'azione migliore da eseguire: ad esempio, recluterà i processori relativi al movimento del braccio in modo da afferrare la palla in movimento.

La **GWT** è una teoria generale e sono state proposte diverse implementazioni. La principale implementazione è LIDA (Learning Intelligent Distributed Agent) sviluppata da Franklin e colleghi

nel corso degli anni [22]. La prima versione di LIDA, denominata IDA [23], è stata utilizzata come sistema di gestione del personale per la Marina degli Stati Uniti. L'obiettivo di IDA era di assegnare i marinai a nuovi compiti al termine del loro turno di servizio. L'interazione con i marinai avveniva tramite e-mail in linguaggio naturale; IDA era in grado di negoziare con i marinai le nuove assegnazioni e scrivere i relativi ordini di servizio.

Un altro approccio alla coscienza artificiale riguarda **la teoria del modello interno**. L'ipotesi del modello interno afferma che un agente, quando agisce in modo cosciente, opera in base a un modello interno di se stesso e del mondo esterno. Il modello interno consente all'agente di simulare le proprie azioni e di valutarne gli esiti prima di compierle nell'ambiente esterno. In questo modo, l'agente può generare aspettative sul corso degli eventi nel mondo e sui risultati delle azioni. Dennett [24] parla di creature “popperiane” in grado di generare teorie sul mondo esterno e di simulare esperimenti nel loro ambiente interno.

Secondo Hesslow [25], **il modello interno permette al cervello di simulare le azioni**, di simulare le percezioni e di generare anticipazioni sugli eventi futuri. Hesslow sostiene che i pensieri coscienti si basano su queste simulazioni. Poiché la simulazione della percezione è legata agli input sensoriali generati internamente che assomigliano alla percezione del mondo esterno, essa sarebbe accompagnata dall'esperienza del modello interno del mondo. In breve, l'ipotesi del modello interno afferma che la coscienza nasce dall'interazione tra il modello interno dell'agente e il modello interno del mondo.

Bongard et al. [26] hanno costruito **un robot “stella marina”** cioè un robot a quattro zampe che genera un modello interno in 3D di se stesso per tentativi ed errori, mediante opportuni algoritmi genetici. Il robot utilizza i propri sensori ed attuatori per dedurre un modello interno del proprio corpo, e poi utilizza il modello del corpo per imparare la locomozione. Il robot è resiliente: in caso di danni, ad esempio una gamba rotta, il robot può generare nuovamente il nuovo modello del suo corpo e imparare la locomozione ancora con il modello del corpo danneggiato. Recentemente, Lipson e colleghi [27] ha esteso questo approccio ad un braccio robotico in grado di costruire un modello di se stesso mediante un sistema di telecamere che inquadra il braccio stesso.

In questo filone, Chella e Macaluso [28] hanno realizzato **“CiceRobot,”** un robot guida museale basato sull'ipotesi del modello interno. Il robot è un caso di studio di diverse capacità associate agli aspetti funzionali della coscienza: la costruzione e il mantenimento del modello interno dell'ambiente e di se stesso; l'attenzione alle entità rilevanti nell'ambiente; l'integrazione delle informazioni provenienti da fonti diverse e da parti diverse della stessa fonte; la generazione di aspettative sui possibili eventi nell'ambiente; auto-monitorarsi; la simulazione di stati emotivi; l'elaborazione di informazioni globalmente disponibili al robot. Il robot è stato inoltre dotato di capacità introspettive; l'introspezione si basa sulla percezione di ordine superiore del robot, nel senso che la percezione di primo ordine del robot è la percezione immediata dell'ambiente, mentre la percezione di ordine superiore è la percezione del mondo interno del robot [29]. L'architettura è stata dimostrata a bordo di un robot che ha condotto numerose visite guidate al Museo archeologico di Agrigento.

Un ulteriore filone riguarda **lo studio sulle relazioni complesse tra emozioni, processi decisionali e coscienza**. In questo filone si fa comunemente riferimento alle teorie di Antonio Damasio [30][31] basate su risultati dalle neuroscienze. Bosse et al. [32] hanno proposto un modello computazionale delle teorie di Damasio, e Cominelli et al. [33], prendendo spunto dalle teorie di Damasio e dal modello di Bosse, hanno realizzato il sistema cognitivo SEAI (Social Emotional Artificial Intelligence) in via di implementazione in un effettivo robot umanoide [34].

Il manifesto sulla coscienza dei robot

Il dibattito sulla coscienza nei robot oggi è molto vivo e vitale. Chella e Manzotti [35] hanno proposto un manifesto sulla coscienza nei robot in cui sono discusse alcune delle principali sfide del campo. La sfida più importante riguarda se la coscienza sia implementabile su un robot oppure se è qualcosa che sfugge ai modelli computazionali ed appartiene soltanto al mondo della biologia. Secondo questa linea di pensiero, soltanto **un essere vivente può essere cosciente e quindi un robot, per quanto sofisticato, non lo sarà mai.**

Un altro dibattito riguarda il ruolo della corporeità: un essere senza corpo come un computer può essere cosciente, oppure è necessario avere un corpo con i sensori quali le telecamere per vedere, e degli attuatori come le braccia e le gambe per potersi muovere, toccare, guardare, e così via? In sostanza, un robot potrà essere cosciente mentre un computer non potrà mai essere cosciente.

Un problema correlato riguarda il ruolo dell'ambiente esterno: **la coscienza risiede nel cervello oppure è qualcosa relativa anche all'ambiente che ci circonda?** Queste problematiche sono importanti perché, riportate all'essere umano, negherebbero la capacità della coscienza al cervello, senza un corpo e senza un ambiente esterno, di essere cosciente.

Inoltre, può un robot provare realmente qualcosa, come un dolore o una gioia, oppure potrà soltanto fingere? Un dibattito correlato riguarda la misura della coscienza: come faremo ad accorgerci se un robot è veramente cosciente? Potrebbe benissimo non essere cosciente ma imitare alla perfezione il fatto di provare qualcosa, ma senza provare realmente nulla.

I test della coscienza nei robot

In letteratura sono stati proposti **numerosi test** per misurare l'eventuale grado di coscienza nei robot: Elamrani e Yampolskiy [36] hanno analizzato e confrontato un gran numero di proposte. Molti test proposti sono variazioni del famoso *test di Turing* [37] dell'imitazione del comportamento umano. Questo test è esposto a tante critiche: è necessaria una persona cosciente che valuta le prestazioni del robot, è basato sul linguaggio ma un robot cosciente potrebbe non avere il linguaggio, e così via. Inoltre, gli ultimi sistemi di generazione automatica del linguaggio come ChatGPT che non sono sicuramente coscienti, possono superare il test grazie alla loro ricchezza espressiva.

In questo filone, **Susan Schneider ha proposto l'Artificial Consciousness Test** [38], una variante del test di Turing in cui le domande da porre sono incentrate sulla qualità dell'esperienza interiore del robot.

Un'altra fonte di ispirazione per i test sulla coscienza è il famoso test dello specchio per i primati [39]. In questo caso, un robot dovrebbe riconoscere e descrivere sé stesso e i suoi movimenti guardando attraverso uno specchio, anche in presenza di altri robot. Alcuni robot, come il robot già citato del gruppo di Lipson e il robot dotato di discorso interiore sviluppato da Pipitone e Chella [40][41] hanno peraltro già superato semplici versioni di questo test.

Conclusioni

L'eventuale avvento di robot coscienti porterebbe a preoccupazioni etiche così come a questioni relative all'integrazione sociale di tali robot, come ha discusso Mark Coeckelbergh [42].

Se un robot fosse capace di soffrire, allora dovrebbe essere trattato come una persona, anche legalmente. Questi argomenti possono costringerci a rivedere le nostre definizioni del concetto di persona. Se affermiamo che un robot è cosciente, allora la responsabilità morale del sistema per le sue azioni deve essere riconosciuta legalmente. D'altra parte, potremmo dover concedere diritti ai robot coscienti, come il diritto di non essere spenti. In sintesi, lo studio della coscienza dei robot è un campo di ricerca che offre grandi opportunità, e problematiche etiche che non possono essere sottovalutate.

Bibliografia

- [1] Whang, O. 2023. 'Consciousness' in Robots Was Once Taboo. Now It's the Last Word. *The New York Times* Jan 6, 2023 <https://www.nytimes.com/2023/01/06/science/robots-artificial-intelligence-consciousness.html>
- [2] Vimal, R. L. P. (2009). Meaning attributed to the term 'consciousness' – an overview. *J. Conscious. Stud.* 16, 9–27.
- [3] Searle, John R. (2000). "Consciousness." *Annual Review of Neuroscience* 23:557–578.
- [4] Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 2, 227 – 247.
- [5] Chalmers, D. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219.
- [6] Nagel, T. (1974). What is like to be a bat? *Philos. Rev.* 83, 435–450.
- [7] Dehaene, S., Lau, H., and Kouider, S. (2017). What is consciousness, and could machines have it? *Science* 358, 486–492.
- [8] Floridi, L. (2005). Consciousness, agents and the knowledge game. *Mind Mach.* 15, 415–444.
- [9] Holland, O. (2003). Robots with internal models – a route to machine consciousness? *J. Conscious. Stud.* 10, 77–109.
- [10] Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* 6, 242–247.
- [11] O'Regan, J. K., and Noë, A. (2001) A sensorimotor account of vision visual consciousness. *Behav. Brain Sci.* 24, 939–973.
- [12] Aleksander, Igor. (1992). Capturing Consciousness in Neural Systems. In *Artificial Neural Networks 2, Proceedings of ICANN 1992 Conference*, 17–22. Amsterdam: Elsevier.
- [13] Kock, Christof. (2001). Final Report of the Workshop *Can a Machine be Conscious*, The Banbury Center, Cold Spring Harbor Laboratory http://www.theswartzfoundation.org/abstracts/2001_summary.asp
- [14] Rushby, John e Sanchez, David. (2018). "Technology and Consciousness Workshops Report." Computer Science Laboratory, SRI International, Menlo Park CA USA

<http://www.csl.sri.com/users/rushby/papers/techconscwks2017.pdf>

[15] Reeke, George N., Olaf Sporns, e Gerald M. Edelman. (1990). “Synthetic Neural Modeling: The ‘Darwin’ Series of Recognition Automata.” *Proceedings of the IEEE* 78 (9): 1498–1530.

[16] Krichmar, J. L., Nitz, D. A., Gally, J. A. and Edelman, G. M. (2005). Characterizing functional hippocampal pathways in a brain-based device as it solves a spatial memory task. *Proceedings of the National Academy of Sciences of the USA* 102 (6): 2111-6.

[17] Oizumi, M., Albantakis, L. and Tononi, G. 2014. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLOS Computational Biology* 10 (5): e1003588.

[18] Koch, C. (2019). *The Feeling of Life Itself: Why Consciousness Is Widespread but Can't Be Computed*. Cambridge, USA: MIT Press.

[19] Aleksander, Igor, and Barry Dunmall. 2003. Axioms and Tests for the Presence of Minimal Consciousness in Agents. *Journal of Consciousness Studies* 10 (4–5): 7–18.

[20] Aleksander, I. 2005. *The World in My Mind, My Mind in the World*. Exeter, UK: Imprint Academic.

[21] Baars, B. J. 1997. *In the Theater of Consciousness. The workspace of the mind*. Oxford, UK: Oxford University Press.

[22] Franklin, S., Madl, T., D’Mello, S., Snider, J. 2014. LIDA: A Systems-level Architecture for Cognition, Emotion, and Learning. *IEEE Transactions on Autonomous Mental Development* 6 (1): 19-41.

[23] Franklin, S. 2003. IDA - A conscious artifact? *Journal of Consciousness Studies* 10 (4-5): 47-66.

[24] Dennett, D. 1996. *Darwin's Dangerous Idea*. NY: Simon & Schuster.

[25] Hesslow, G. 2002. Conscious Thought as Simulation of Behaviour and Perception. *Trends in Cognitive Sciences* 6 (6): 242-247.

[26] Bongard, J., Zykov, V. and Lipson, H. 2006. Resilient Machines Through Continuous Self-Modeling. *Science* 314: 1118-1121.

[27] Chen, B., Kwiatkowski, R., Vondrick, C, Lipson, H. 2023. Fully body visual self-modeling of robot morphologies. *Sci. Robot.*, 7 (68), eabn1944

[28] Chella, A., Macaluso, I. 2009. The Perception Loop in Cicerobot, a Museum Guide Robot. *Neurocomputing* 72:760–766.

[29] Chella, A., Frixione, M., Gaglio, S. 2008. A Cognitive Architecture for Robot Self-Consciousness. *Artificial Intelligence in Medicine* 44:147–154.

[30] Damasio, A. (1994). *Descartes’ Error: Emotion, Reason, and the Human Brain*. New York: Putnam.

[31] Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt.

[32] Bosse, T., Jonker, C. M., Treur, J. (2008). Formalisation of Damasio’s theory of emotion, feeling and core consciousness. *Conscious. Cogn.* 17, 94–113

[33] Cominelli L, Mazzei D and De Rossi DE (2018) SEAI: Social Emotional Artificial Intelligence Based on Damasio's Theory of Mind. *Front. Robot. AI* 5:6. doi: 10.3389/frobt.2018.00006

[34] Cominelli, L., Hoegen, G., De Rossi, D. (2021) Abel: integrating humanoid body, emotions and time perception to investigate social interaction and human cognition *Appl. Sci.* 11 1070

[35] Chella, A., Manzotti, R. 2009. Machine Consciousness: A Manifesto for Robotics. *International Journal of Machine Consciousness* 1 (1): 33–51.

[36] Elamrani, A., Yampolskiy, R.V. 2019. Reviewing Tests for Machine Consciousness. *Journal of Consciousness Studies*, 26, 5–6, 35–64.

[37] Turing, Alan, M. 1950. "Computing Machinery and Intelligence." *Mind*, LIX, 236, 433–460.

[38] Schneider, Susan, e Edwin Turner. 2017. Is Anyone Home? A Way to Find Out if AI Has Become Self-Aware. *Scientific American* (blog). <https://blogs.scientificamerican.com/observations/is-anyone-home-a-way-to-find-out-if-ai-has-become-self-aware/>

[39] Gallup Jr., G.G. (1970) Chimpanzees: Self-recognition, *Science* 167(3914), pp. 86–87.

[40] Pipitone, A., Chella, A. 2021. Robot passes the mirror test by inner speech. *Robotics and Autonomous Systems* 144, 103838

[41] Chella, A., Pipitone, A., Morin, A., Racy, F. 2020. Developing Self-Awareness in Robots via Inner Speech. *Frontiers in Robotics and AI* 7:16.

[42] Coeckelbergh, Mark. 2022. Robot Ethics. MIT Press.

Giocando s'impara? Alla ricerca delle proprietà educative “elevate” di giochi e videogiochi

La letteratura scientifica ha esplorato le opportunità di giochi e videogiochi per l'educazione e il cambiamento psicologico, ma si è focalizzata principalmente sugli aspetti più semplici. Al contrario, evidenze recenti sfidano queste concezioni, concentrando l'attenzione sulle proprietà di fiction e immedesimazione

Di **Stefano Triberti**, professore associato di Psicologia Generale presso Università Telematica Pegaso

Maria Concetta Carruba, ricercatore in Pedagogia Speciale presso Università Telematica Pegaso

Da più di trent'anni la **ricerca psicologica** esplora le **proprietà di giochi e videogiochi** per il potenziamento delle **abilità cognitive**. È indubbio che i giochi in generale, in quanto strutture di regole da padroneggiare e applicare per ottenere un determinato scopo, costituiscono **stimoli complessi** per la nostra mente, capaci di coinvolgerci a livello di ragionamento, memoria, sistemi motivazionali e risposta motoria.

Il ruolo del gioco nei processi di apprendimento

L'origine dell'idea che i *giochi* in generale possano allenare e sviluppare le abilità cognitive, e quindi **“renderci più intelligenti”**, si perde nei tempi antichi: si può discutere del fatto che già Platone, paragonando il gioco all'atto creativo degli dei e alla natura del mondo (come già accennava Eraclito), riconoscesse all'azione di giocare una funzione positiva per la crescita e la maturazione.

Anche la letteratura scientifica in ambito pedagogico conferma **il ruolo centrale del gioco nei processi di apprendimento** tanto in quello spontaneo che in quello intenzionale e strutturato. Gli autori classici avevano già evidenziato nei loro lavori gli aspetti educativi del gioco, l'importanza della metodologia attiva nei processi didattici o, ancora, vere e proprie classificazioni del gioco (Piaget 1951; Vygotskij 1967; Bruner, Jolly, Sylva 1976) utili per coglierne le potenzialità da un punto di vista didattico: rendere la didattica attiva consente allo studente un apprendimento più esperienziale. In tempi più recenti, però, anche a fronte delle molteplici opportunità e produzioni digitali, è possibile rintracciare nei giochi e nei videogiochi il ruolo di veri e propri **“facilitatori”** per l'apprendimento che usano il gioco come sfondo e permettono all'allievo un apprendimento trasversale, indiretto (Almon, 2003; Samuelsson, Carlsson, 2008; Bodrova, Germeroth, Leong, 2013; Bateson, Martin, 2013). Lo studente “apprende”, dal latino *apprehendere* che significa “ricevere e trattenere nella mente”, subendo meno il carico cognitivo dei processi di apprendimento tradizionali e, al contempo, diviene co-costruttore del proprio sapere. Prospettiva che evidenzia in quale misura il gioco permetta di sviluppare, allenare, interiorizzare e consolidare alcune caratteristiche e abilità che, pur se non specificamente coinvolti nel processo di apprendimento, ne

risultano un sostegno al processo: **l'autoregolazione, la flessibilità, la negoziazione, il decision making, il problem solving, la creatività** (Gily, 2002; Lazzeri & Bocci, 2022).

Giocare diviene, dunque, l'occasione per giocare con gli apprendimenti, per mettersi in gioco. Pratica consolidata e pedagogicamente riconosciuta per il bambino che, nel tempo viene spesso tralasciata o, ancor peggio, usata male. **Il gioco, invece, rappresenta non solo uno strumento ma un vero e proprio “spazio” di apprendimento** entro il quale si entra in contatto con la motivazione personale, la capacità di leggere il “contesto” e accoglierne le regole (netiquette), la propensione all'esplorazione, all'autocontrollo e alla socialità. Senza limite alcuno legato all'età.

Anche relativamente al gioco digitale e al videogioco, nell'Agenda 2030 (UN, Horizon 2030, 2015) si sottolinea l'importanza di allenare le competenze digitali già a partire dalla scuola dell'Infanzia.

I giochi potenziano le abilità cognitive?

È possibile comunque sostenere che tali strumenti ludici abbiano effetti dimostrabili sulle capacità cognitive? Per quanto riguarda i **videogiochi**, è importante ricordare gli studi pionieristici di Green e Bavelier (2003; 2006; 2012; Achtman, Green & Bavelier, 2008): pubblicarono diversi studi dai primi anni Duemila da cui sembrava evincersi che i giocatori di videogiochi sviluppavano diverse **abilità cognitive** (come la velocità di apprendimento o l'attenzione divisa) a un livello superiore rispetto ai non-giocatori. I risultati di queste ricerche sono stati in parte **ridimensionati** dalla letteratura successiva (Murphy & Spencer, 2009; Trisolini, Petilli & Daini, 2018): Green e Bavelier, infatti, non misuravano gli effetti di sessioni di gioco, bensì confrontavano giocatori assidui e non giocatori in termini di performance a test cognitivi tradizionali.

È stato osservato che tale approccio metodologico non consentiva di considerare **l'intervento delle variabili motivazionali**. È stato anche suggerito da studi più approfonditi (Irons, Remington & McLean, 2011) che i videogiocatori potrebbero essere capaci di adattare la loro **performance** alle richieste del compito nel corso dello stesso (una sorta di capacità tattica, certamente richiesta sovente dai videogiochi), ma non per questo possano essere considerati “superiori” in termini di abilità attentive e cognitive.

Gli aspetti positivi trasversali del gioco

Studi recenti permettono di correlare l'esposizione a nuovi strumenti di apprendimento, quali giochi e videogiochi appunto, a **conformazioni o strutture mentali ben diverse dalle generazioni precedenti** che non hanno avuto modo di relazionarsi con strumenti e opportunità di questo tipo. Oggi si parla addirittura di Neurodidattica (Rivoltella, 2012) e di abilità e competenze che sembrano mutando in riferimento alle pratiche d'uso di device e all'esposizione ai videogiochi. Latham, Patson e Tippet (2013) nei loro lavori dimostrano che diversi decenni di videogiochi hanno certamente inciso rispetto ad alcune abilità cognitive. Tra le principali evidenze degli esiti di queste loro ricerche possiamo citare:

1. **una migliore coordinazione oculo manuale** che peraltro non sembrerebbe dipendere dall'intensità d'uso;
2. **tempi di reazione notevolmente più brevi;**
3. **miglioramento dell'elaborazione visuo-spaziale;**
4. **più ampia capacità di anticipazione visiva** e di strategie di ricerche visive;
5. **migliore gestione dell'attenzione in situazioni dinamiche.**

Accanto a questi aspetti positivi trasversali, non sono mancati poi studi che ci permettono di affermare che, se progettati e usati in modo sistematico, **giochi e videogiochi possono rappresentare un facilitatore per l'apprendimento** anche per gli studenti con profili di funzionamento non neurotipici. Negli studi di Franceschini et al. (2013) ma anche in quelli di Sibilio, Fulgione, Di Tore (2014) emerge come un training con videogiochi d'azione ha portato a un miglioramento delle capacità di lettura per il campione di studenti con dislessia preso in esame.

Gli studi sul ruolo del gioco nel processo di apprendimento nei contesti formali

Dal punto di vista **pedagogico**, dobbiamo certamente a Papert & Harel (1991) prima e a Resnick (1995) poi i primi importanti lavori sul ruolo del gioco nel processo di apprendimento nei contesti formali. Gli autori definiscono giochi e videogiochi come una vera e propria palestra per il “pensiero computazionale”. Semplificando, in questo contesto il pensiero computazionale può essere inteso come un complesso processo che permette di **scomporre un macro “compito” in più sotto compiti** per meglio leggere e comprendere lo stesso allenando processi mentali elevati quali il ragionamento, l'immaginazione, le inferenze e le rappresentazioni (Giroto, Legrenzi, 1999; Ferranti, Agosto, 2016). Così definito, il pensiero computazionale ha assunto un vero e proprio valore pedagogico e riconosciuto come fondamentale per scomporre ma anche per decodificare e, conseguentemente, programmare azioni e procedure ripetibili (algoritmi riutilizzabili).

Questo tipo di pensiero associato al gioco richiama alla **capacità dei progettisti dell'educazione (gli insegnanti) di proporre attività didattiche e di apprendimento che permettano all'allievo di sperimentarsi con attività piacevoli ma sfidanti** perché possano attivare quei processi mentali necessari per padroneggiare situazioni complesse (Papert, 1988). Per molto tempo il gioco didattico, propriamente inteso, è stato rappresentato da software didattici chiusi (che non consentono modifiche e personalizzazioni). Il software didattico chiuso è concepito con un sistema predefinito (chiuso, appunto) nel quale il giocatore impara attraverso un susseguirsi di attività lineari. Nascono con scopi didattici espliciti e ruotano intorno a una disciplina o un contenuto specifico della stessa. Un incentivo per “imparare bene e velocemente”, volendo usare le parole di Papert. I giochi e videogiochi di ultima generazione sono, invece, addirittura “aperti” e consentono un vero e proprio apprendimento per scoperta lasciando al giocatore la possibilità di orientare il processo.

Le possibilità dei videogiochi di potenziare abilità cognitivo-sociali

Parallelamente, la letteratura ha esplorato le possibilità dei videogiochi di potenziare abilità cognitivo-sociali, come ad esempio la regolazione emotiva. Il potenziamento di queste abilità è strettamente legato, nelle prime formulazioni di questi studi, alla componente motivazionale: si evidenzia che i videogiochi vengono spesso giocati per distrarsi dalle emozioni e dalle esperienze negative, combattere lo stress, rifugiarsi in un “posto sicuro” emotivo (*escapism*), in altre parole a fini di *mood repair* (Bowman & Tamborini, 2012; Reinecke et al., 2012; Rieger et al., 2014). Alcuni studi successivi utilizzano metodologie più sofisticate ed esplorano la possibilità di *creare* videogiochi appositamente per indurre frustrazione e poi insegnare strategie per gestirla (Rodriguez et al., 2015; Vara et al., 2016), oppure per rappresentare il processo di regolazione delle emozioni e così allenarlo (Scholten et al., 2016; Cejudo & LaTorre, 2015).

Queste e altre evidenze hanno portato allo sviluppo di **concetti particolarmente fortunati come quelli di edutainment, gamification, serious game, game for social change** (Triberti & Argenton, 2013; Maestri, Polsinelli & Sassoon, 2018; Kappa, 2012): una vera e propria cultura del gioco e del videogioco che punta coscientemente a sfruttarne le opportunità formative e didattiche già al livello dello sviluppo e del game design. Diventa piuttosto comune incontrare progetti di ricerca e di

intervento che contengono, come elemento secondario oppure come vera protagonista, una attività di gioco o videogiochi volta ad allenare abilità cognitive, fisiche, emotivo-sociali, il tutto radicato nella credenza che tali strumenti dispongano di una particolare efficacia nel promuovere il cambiamento.

È possibile oggi confermare questa credenza? Solo in parte; review sistematiche e meta-analisi focalizzate sui serious game in diversi ambiti (Girard, Ecalle, & Magnan, 2013; Ypsilanti et al., 2014) riportano evidenze non completamente conclusive. I problemi della letteratura riguardano in particolare **misurazioni dell'efficacia non chiare o incommensurabili**, e anche una eccessiva diversità tra le tipologie di gioco/videogioco utilizzate nei diversi studi. Il risultato più frequente è che i prodotti ludici si distinguono consistentemente da altri strumenti di intervento in termini di motivazione positiva e coinvolgimento, ma non è possibile dire con chiarezza se siano universalmente più efficaci nel raggiungere gli obiettivi formativi.

Cambiare il nostro sguardo sui giochi

È possibile sostenere che sia necessario **complessificare la rappresentazione che la letteratura scientifica ha di giochi e videogiochi**. Da una parte, review successive focalizzate sull'utilizzo dei videogiochi per la promozione della regolazione emotiva (Villani et al., 2018) trovavano risultati interessanti: l'evidenza di studi multipli sembrava suggerire che maggiori opportunità per l'esercizio di abilità emotive si potevano trovare nel **gioco assiduo con prodotti commerciali** piuttosto che nell'utilizzo di serious game. Gli autori speculavano che ciò fosse dovuto alla possibilità per i giocatori di sperimentarsi non solo con le esperienze di “mera interazione” spesso supportate da videogiochi creati *ad hoc* (per esempio: attività semplici come far scoppiare bolle, costruire strutture, *platforming*, ecc.), ma anche e soprattutto di potersi immergere in storie complesse e sistemi di relazioni. In questo senso le proprietà più potenti di giochi e videogiochi per mettere alla prova abilità emotive e sociali sarebbero le proprietà di *fiction* nelle quali il giocatore può immergersi per consistenti periodi di tempo: pur non volendo con questo suonare dispregiativi, il semplice “giochino” con vaghe proprietà simboliche o pensato principalmente per distrarre o rilassare non promuove il medesimo cambiamento che può ottenere una complessa storia nella quale il giocatore può fare esperienza di forti emozioni, scelte morali, relazioni complesse. Contesti simili agiscono come veri e propri “laboratori” per l'esercizio di abilità sociali (Triberti, Villani & Riva, 2015; Triberti, Di Pasquale e Riva, 2018).

Un secondo elemento rilevante, che a nostro avviso merita interesse di ricerca per una concezione rinnovata dei giochi e del loro potenziale per il cambiamento, sono alcune **tendenze nell'evoluzione dei giochi stessi**, per come sono rilevabili nella società contemporanea, al livello tanto delle proposte commerciali attuali quanto dei mutamenti della cultura *pop* che li circonda. Fino a qualche anno fa si discuteva dell'emergere di giochi “semplici” e “user-friendly”; vi è stata senz'altro una importante deviazione nel mondo di giochi e videogiochi, volta a catturare l'attenzione di gruppi demografici differenti da giovani e adolescenti. Emerse a livello globale il concetto di “**casual gaming**” (Wilson & Leaver, 2016), giochi rivolti a famiglie, adulti e anziani che tendenzialmente hanno poco tempo e interesse per sviluppare competenze e abilità complesse: così i **party game**, numerosi giochi per piattaforme *mobile* o console rivolte a famiglie, furono sviluppati per facilitare un ingresso immediato nell'ambiente di gioco e privilegiare relax e divertimento sugli elementi di complessità.

Se da una parte questa tendenza è ancora presente e rappresenta una importante fetta del mercato di giochi e videogiochi, è possibile a nostro avviso notare anche **un “ritorno di massa” di prodotti rivolti a giocatori che cercano al contrario sfide difficili e coinvolgenti a livello cognitivo/intellettuale**. Rappresentanti di questa tendenza si possono rilevare per esempio:

- **nel potente ritorno dei libri game**, caratterizzati da meccaniche di gioco ed enigmi innovativi e complessi, rinnovamento di cui tra l'altro l'Italia è considerabile una importante locomotiva, tra nuovi autori e case editrici dedicate (Giove, 2022);
- nell'emergere e nella diffusione negli ultimi anni delle *escape rooms*, esperienze interattive e di gruppo di grande successo, che riportano l'esperienza di gioco al di fuori del digitale e delle case degli utenti (Veldkamp et al., 2020; Wiemker, Elumir, & Clare);
- nell'interesse degli sviluppatori e dei giocatori verso **giochi e videogiochi di tipo analitico, riflessivo, investigativo** (per fare un esempio, sulla piattaforma Steam che supporta l'acquisto di una vasta proporzione dei giochi esistenti per computer, giochi con elementi di "mistero" e "rompicapo" superano oggi di gran lunga i più noti "sparatutto" e "picchiaduro");
- nel **"return of the Whodunnit"** di cui scrivono studiosi e giornalisti di media, ovvero il rinnovato interesse, tra diversi media come videogiochi, podcast e cinema, per i gialli tradizionali in cui il *focus* narrativo è sul processo di investigazione e sulla complessità degli enigmi.

Conclusioni

Che rilevanza ha questo per la ricerca psicologica e pedagogica? È possibile che alcune opportunità dei giochi e dei videogiochi siano state finora, almeno in parte, sotto-sfruttate. Se oggi è dimostrato che sistemi interattivi divertenti supportano la motivazione e l'impegno, e migliorano abilità cognitive di base, è possibile immaginare che lo sfruttamento delle proprietà più "alte" di giochi e videogiochi (come le proprietà di fiction e immedesimazione, o l'inserimento di enigmi e problemi realmente complessi e sfidanti) presentino ancora delle sorprese. La ricerca psicologica può puntare innanzitutto alla *profilazione* degli utenti di giochi e videogiochi di generi peculiari, meno noti o emersi solo recentemente al livello degli interessi dell'utenza, cercando di comprendere le motivazioni che supportano un impegno notevole o un rilevante dispendio di "energie" a livello cognitivo, e allo stesso tempo può esplorare i benefici attesi o effettivi che tali giochi e videogiochi hanno sulla vita e le capacità dei fruitori.

Evidenze di questo tipo offrirebbero **nuovi spunti anche alla ricerca pedagogica**, soprattutto nell'ottica di impostare la didattica di concetti complessi oppure supportare il cambiamento di atteggiamenti e comportamenti a livello morale ed esistenziale, permettendo agli utenti della formazione di mettere alla prova sé stessi all'interno di scenari simulati e tuttavia immersivi. La nostra ipotesi è che l'utilizzo di giochi di questo tipo (proprietà di fiction elevate e/o complesse richieste cognitive) possa non solo promuovere cambiamenti comportamentali e di apprendimento/educazione, ma anche garantire che tali cambiamenti perdurino nel tempo e contribuiscano a formare caratteristiche e capacità stabili nelle persone.

Infatti, dal punto di vista pedagogico e didattico il gioco e il videogioco non rappresentano solo l'occasione per una proposta didattica attiva e interattiva e neppure soltanto l'occasione per promuovere un apprendimento più divertente e piacevole per l'allievo; essi rappresentano una "palestra progettuale" anche per l'insegnante che può prendere spunto dai principi di apprendimento su cui si fondano giochi e videogiochi per applicare gli stessi anche in altri contesti. Gee (2003) individua ben 36 principi tra i quali possiamo evidenziare i quattro più significativi in termini didattici:

1. **I giochi ben fatti limitano la decontestualizzazione e il sovraccarico informativo e forniscono informazioni su richiesta in tempo reale** relativamente al contesto d'uso e connesse agli obiettivi didattici e/o educativi esplicitati;

2. **I giochi ben fatti propongono all'allievo sfide graduali e in crescendo:** sono sfidanti a tal punto da motivare al successo ma non eccessivamente frustranti per evitare di demotivare il fruitore con sfide non in linea con il livello di competenza raggiunto;
3. **I videogiochi intesi come “software aperti”, consentono al fruitore di essere al contempo giocatore e costruttore delle sfide ludiche** (le scelte operate dal giocatore variano il contesto e lo rendono strettamente personale l'esperienza di gioco);
4. **nei videogiochi ben fatti ci si misura con sfide che vengono praticate fino a quando non si verifica una nuova padronanza** di ordine superiore offrendo al giocatore la possibilità di effettuare generalizzazioni che divengono un vero e proprio “ciclo di esperienza”.

In accordo con questa visione, proponiamo quindi una rinnovata attenzione a giochi e videogiochi che sappia riconoscere le loro evoluzioni e la loro complessità, allo scopo di esplorare ulteriormente le opportunità offerte da questi complessi media per la crescita delle persone e per la promozione della salute e del benessere.

Bibliografia

Achtman, R. L., Green, C. S., & Bavelier, D. (2008). Video games as a tool to train visual skills. *Restorative neurology and neuroscience*, 26(4-5), 435-446.

Almon J., (2003). The vital Role of Play in Early Childhood Education, in, S. Olfman (ed), All work and no Play... How Educational Reforms are Harming our Preschoolers. Praeger Publishers/GreenWood Publishing Group, Westport (CT), pp. 17-42.

Bateson P., Martin P., (2013), Play, Playfulness, Creativity and Innovations,, Cambridge University Press, Cambridge.

Bodrova E., Germeroth C., Leong D.J., (2013), Play and Self-Regulation: Lessons from Vygotsky, in, American Journal of Play, 6, pp. 111-23.

Bowman, N. D., & Tamborini, R. (2012). Task demand and mood repair: The intervention potential of computer games. *New Media & Society*, 14(8), 1339-1357.

Bruner J.S., Jolly A., Sylva K., (1976). Play: its role in the development and evolution, Harmondsworth, Penguin.

Cejudo, J., & Latorre, S. (2015). Effects of the Spock videogame on improving emotional intelligence in adolescents. *Electronic Journal of Research in Educational Psychology*.

Nazioni Unite (2015). Trasformare il nostro mondo: l'Agenda 2030 per lo Sviluppo Sostenibile. *Risoluzione adottata dall'Assemblea Generale il*, 25.

Di Tore S., Fulgione M., Sibilio M., (2014), Dislessia e videogames. Il potenziale didattico dei videogiochi, in, Mediterranean Journal of Social Sciences, 5,23, pp. 1165.

Ferranti C., Agoto V., (2016), Tecnologie e pensiero computazionale: proposte per il curricolo di tecnologia nella scuola primaria, in, Didamatica 2016.

Franceschini S. Gori S., Ruffino M., Viola S., Molteni M., Facchetti A., (2013), Action Video Game Make Dyslexic Children Read Better, in *Current Biology*, 23, 6, pp. 462-6.

- Gily C., (2002), *In-lusio, il gioco come educazione estetica*, Garus, Napoli.
- Giove, C. (2022) *Prefazione*. In AA. VV., *Creature, Gate on Games - Zerosem S.R.L.*, Guamo (Lucca).
- Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of computer assisted learning*, 29(3), 207-219.
- Giroto V., Legrenzi P., (1999), *Psicologia del Pensiero*, Il Mulino, Bologna.
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, 423(6939), 534-537.
- Green, C. S., & Bavelier, D. (2006). The cognitive neuroscience of video games. *Digital media: Transformations in human communication*, 1(1), 211-223.
- Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current biology*, 22(6), R197-R206.
- Irons, J. L., Remington, R. W., & McLean, J. P. (2011). Not so fast: Rethinking the effects of action video games on attentional capacity. *Australian Journal of Psychology*, 63(4), 224-231.
- Kappa, K. M. (2012) *The gamification of learning and instruction*. Wiley
- Latham A.J., Patson L.L., Tippet L.J., (2013) *The Virtual Brain: 30 Years of Video Game Play and Cognitive Abilities*, in, *Frontiers in Psychology*, 4, 629.
- Lazzeri, M., & Bocci, F. (2022) *Giochi e videogiochi fanno bene alla salute? Modalità ed esempi*. *I Quaderni di Agenda Digitale*, 10
- Maestri, A., Polsinelli, P., & Sassoon, J. (2018) *Giochi da prendere sul serio. Gamification, storytelling e game design per progetti innovativi*. Franco Angeli
- Murphy, K., & Spencer, A. (2009). Playing video games does not make for better visual attention skills. *Journal of Articles in Support of the Null Hypothesis*, 6(1).
- Papert S., Harel I., (1991), *Situating Constructionism*, in *Costructionism*, 36, 2.
- Piaget, J. (1951). *Play, dreams and imitation in childhood*. Routledge, London.
- Reinecke, L., Tamborini, R., Grizzard, M., Lewis, R., Eden, A., & Bowman, N. (2012). Characterizing behavioral affinity as needs satisfaction: Predicting selective exposure to video games and resultant mood repair. *Default journal*, 437-453.
- Resnik M., (2015), *New Paradigms for Computing, New Paradigms for Thinking*, in, A.A. di Sessa, C. Hoyles, R. Noss, L.D. Edwards (eds), *Computers and Exploratory Learning*, vol. 146, Springer-Verlag, Berlin-Heidelbergh, pp. 31-43.

Rieger, D., Wulf, T., Kneer, J., Frischlich, L., & Bente, G. (2014). The winner takes it all: The effect of in-game success and need satisfaction on mood repair and enjoyment. *Computers in Human Behavior*, 39, 281-286.

Rivoltella P.C., (2012), Neurodidattica. Insegnare al cervello che apprende, Raffaello Cortina, Milano.

Rodriguez, A., Rey, B., Vara, M. D., Wrzesien, M., Alcaniz, M., Banos, R. M., & Perez-Lopez, D. (2015). A VR-based serious game for studying emotional regulation in adolescents. *IEEE Computer graphics and applications*, 35(1), 65-73.

Samuelsson I.P., Carlsson M.A., (2008), The Playing Learning Child: Towards a Pedagogy of Early Childhood, in, *Scandinavian Journal of Educational Research*, 52, 6, pp. 623-41.

Scholten, H., Malmberg, M., Lobel, A., Engels, R. C., & Granic, I. (2016). A randomized controlled trial to test the effectiveness of an immersive 3D video game for anxiety prevention among adolescents. *PLoS one*, 11(1), e0147763.

Triberti, S., & Argenton, L. (2013) *Psicologia dei videogiochi: Come i mondi virtuali influenzano mente e comportamento*. Sant'Arcangelo di Romagna: Maggioli

Triberti, S., Di Pasquale, C., & Riva, G. (2018) *I Mondi Virtuali e i Criteri Multipli: Il Problema Morale nei Videogiochi Massivi*. *Lessico di Etica Pubblica*, 1

Triberti, S., Villani, D., & Riva, G. (2015). Moral positioning in video games and its relation with dispositional traits: The emergence of a social dimension. *Computers in Human Behavior*, 50, 1-8.

Trisolini, D. C., Petilli, M. A., & Daini, R. (2018). Is action video gaming related to sustained attention of adolescents?. *Quarterly Journal of Experimental Psychology*, 71(5), 1033-1039.

Vara, M. D., Baños, R. M., Rasal, P., Rodríguez, A., Rey, B., Wrzesien, M., & Alcañiz, M. (2016). A game for emotional regulation in adolescents: The (body) interface device matters. *Computers in Human Behavior*, 57, 267-273.

Villani, D., Carissoli, C., Triberti, S., Marchetti, A., Gilli, G., & Riva, G. (2018). Videogames for emotion regulation: a systematic review. *Games for health journal*, 7(2), 85-99.

Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Soviet psychology*, 5(3), 6-18.

Ypsilanti, A., Vivas, A. B., Räisänen, T., Viitala, M., Ijäs, T., & Ropes, D. (2014). Are serious video games something more than a game? A review on the effectiveness of serious games to facilitate intergenerational learning. *Education and Information Technologies*, 19, 515-529.

I limiti e i rischi di ChatGPT non sono quelli che crediamo, possiamo ancora evitare il peggio

ChatGPT ha riavviato il dibattito sulle Intelligenze artificiali spostando l'attenzione sui loro pregi e difetti e distogliendola dal pericolo che si corre accentrando il potere tecnologico nelle mani di poche aziende

Di **Giuseppe Attardi**, professore di Informatica all'Università di Pisa

ChatGPT ha portato all'attenzione del vasto pubblico la tecnologia dei **Large Language Model (LLM)**, che ne sono alla base. La sua [capacità di interagire e di rispondere](#) in linguaggio naturale in più lingue ha fatto insorgere curiosità e **ha lasciato intendere che possa esibire capacità paragonabili** all'intelligenza umana. Tramite esperimenti scientifici controllati, è stato possibile verificare che questo può succedere, ma solo in compiti di trattamento di testi molto specifici, mentre i Llm sono ben lontani dalle capacità della cosiddetta Artificial General Intelligence (Agi).

Ma quali scenari si prefigurano per un futuro in cui questa tecnologia e le sue evoluzioni rimarranno nappannaggio di poche grandi imprese? E, soprattutto, possiamo ancora fare in modo che questo non accada?

I chatbot e ChatGPT

Finora, la tecnologia principale per costruire i chatbot era basata su **schemi di dialogo preconfezionati** (template) con domande tipiche per ciascuno scopo (intent) con corrispondenti risposte (fulfillment) che i sistemi si limitavano ad adattare, riempiendone i campi con valori estratti dai dialoghi.

Un assistente online, per esempio, poteva essere predisposto a rispondere domande su come accedere a un servizio, come creare un account, come ottenere una password, o come cambiarla in caso di smarrimento. A tutti sarà capitato di imbattersi in tali tipi di chatbot, rimanendo frustrati dall'incapacità di ottenere aiuto al di fuori degli argomenti prefissati.

Amazon, alla ricerca di idee su come migliorare la capacità di dialogo di Alexa, il suo voice assistant, [organizza la Alexa Prize](#), una sfida tra tutte le università del mondo, per "inventing next-generation conversational AI experiences", che ha assegnato un premio di \$500.000 al team vincitore, di cui fa parte Federico Rossetto, un laureato alla Magistrale di Intelligenza Artificiale a Pisa.

ChatGPT Sprints to One Million Users

Time it took for selected online services to reach one million users



* one million backers ** one million nights booked *** one million downloads
Source: Company announcements via Business Insider/LinkedIn



statista 

Perché ChatGPT è diverso

ChatGPT supera i limiti dei chatbot tradizionali combinando tre tecniche:

- un **Large Language Model** (GPT-3.5) di cui sfrutta la capacità di capire e generare frasi in linguaggio naturale
- la messa a punto (**fine-tuning**) sul compito specifico di rispondere a domande
- il **Reinforcement Learning per** imparare a scegliere la mossa migliore di una strategia per raggiungere un obiettivo, in questo caso la risposta migliore, per compiacere l'interlocutore.

I [Large Language Model](#) (Llm) sono modelli di reti neurali profonde ([Deep Learning](#)) in grado di **acquisire una conoscenza dettagliata di una lingua**, ricavandola da enormi quantità di testi, tratti principalmente dal Web.

Per esempio, su Wikipedia si può leggere che Roma è la capitale dell'Italia e Parigi quella della Francia. Si intuisce che un Llm sappia completare la frase “La capitale è _” con la parola “Parigi”, a fronte della domanda “Qual è la capitale della Francia?”: il fine-tuning gli ha insegnato la forma della risposta e l'attention, di cui parleremo dopo, a tenere conto della parola “Francia” dal contesto della domanda.

Ma le possibilità dei Llm **si sono presto dimostrate ben superiori alla loro intrinseca capacità** di completare una frase o di comporre intere storie a partire da un breve spunto iniziale.

Il Reinforcement Learning **utilizza un premio da assegnare al sistema quando la mossa che sceglie è utile** a raggiungere l'obiettivo. Nel caso di ChatGPT l'obiettivo è di soddisfare le richieste dell'interlocutore, e il premio si basa sul confronto tra più risposte possibili.

OpenAI ha raccolto tantissime risposte alternative, ricavate da dialoghi con allenatori umani che interagivano con il chatbot da addestrare e davano un punteggio a quelle migliori. Il Llm di base (GPT-3.5) è stato messo a punto (fine-tuned) per generare la risposta migliore coerentemente a questi esempi di risposte.

Il perché della popolarità di ChatGPT

ChatGPT è diventato popolare perché OpenAI ha messo a disposizione una demo [online per dialogare col chatbot](#) nella propria lingua, anziché doverlo programmare come gli altri Llm.

Milioni di persone lo hanno voluto mettere alla prova e i commenti si sono divisi tra gli entusiasti e i detrattori. I primi erano stupiti e orgogliosi di vedere una piccola creatura alzarsi in piedi e compiere i primi passi, considerandolo un momento cruciale del suo sviluppo. Gli altri si sono sforzati di farla cadere con uno sgambetto o di farla cadere dalla bicicletta, che non aveva mai provato. Cercare domande sbagliate a cui ChatGPT dà risposte sbagliate **è diventato uno sport diffuso**, anziché cercare le domande giuste a cui dà la risposta giusta.

Per farsi un'opinione scientificamente valida, non bastano singoli esempi scelti appositamente, ma occorre innanzitutto capire la tecnologia e i suoi limiti per saperla sfruttare al meglio. Anche coi motori di ricerca, ci siamo rapidamente adattati ai loro limiti: sapendo che si basano sul confronto tra parole chiave della ricerca e parole presenti nei testi, abbiamo imparato a scegliere le parole giuste e a cambiarle quando non ottenevamo i risultati che ci aspettavamo.

In realtà, ChatGPT era stato [preceduto da Galactica](#) di Meta, che svolgeva compiti simili di produrre brevi saggi a tema o riassunti di articoli a partire da una raccolta di testi scientifici. Ma **Galactica fu ritirato dalla circolazione** tre giorni dopo il debutto per una valanga di critiche sui suoi errori e su possibili abusi. ChatGPT ha introdotto una serie di contromisure per proteggersi dagli abusi, ma soprattutto ha avuto la sfrontatezza di mettere alla prova del pubblico un sistema sperimentale, che un'azienda consolidata con miliardi di utenti come Meta non poteva permettersi.

ChatGPT è solo uno dei tanti modi che si stanno scoprendo su **come usare i Llm per una varietà di compiti**. La ricerca sta facendo rapidissimi progressi in questo settore e nuovi modelli vengono sviluppati in continuazione. **Non dobbiamo quindi pensare che ChatGPT sia il meglio** che la tecnologia possa offrire, ma solo un passo di uno sviluppo che continuerà a stupirci.

I Large Language Model

I Llm costituiscono il secondo dei tre inaspettati breakthrough scientifici del Deep learning applicato al [Natural Language Processing](#), avvenuti nel breve periodo di dieci anni.

Il primo breakthrough fu l'invenzione di un metodo per **rappresentare il senso delle parole** con tecniche di apprendimento non supervisionato (self-supervised), ossia bastava fornire a una rete neurale un elevato numero di frasi affinché imparasse a cogliere somiglianze di significato tra le parole che le componevano.

Ogni parola viene rappresentata da un word-embedding, un vettore di centinaia di numeri, ciascuno che in qualche modo **coglie una particolare sfumatura di significato**. Parole di significato simile si trovano vicine tra loro in questo spazio, per esempio Francia, Italia e Germania [sono vicine](#) (basta provare a inserire la parola "France" nel camp di ricerca), facendo supporre che possano essere accomunate da qualcosa che noi chiameremmo la categoria nazione.

Microsoft, Google e Apple saranno altrettanto vicine, **legate forse dal concetto di azienda digitale**.

Categorie e concetti emergono naturalmente, come parole presenti in un certo intorno dello spazio degli embedding, anche più articolati e numerosi dei concetti che si possono ritrovare in dizionari o ontologie curate a mano. **Apple però è un termine ambiguo**, il cui significato dipende dal contesto.

Su questo interviene il secondo breakthrough, con l'introduzione di un meccanismo di attenzione, descritto nell'articolo seminale "Attention is All You Need" (Vaswani, et al., 2017). **Con l'attention si riescono a cogliere legami e relazioni tra le parole** in un contesto e costruire i cosiddetti Transformer, ossia modelli che trasformano una sequenza di input in una sequenza di output, conservando le relazioni tra le parole.

Più in generale, si tratta [di reti neurali](#) utilizzate per elaborare dati sequenziali (quindi frasi, voce, fenomeni con andamenti temporali, eccetera) **che ne elaborano tutti gli elementi in parallelo**, per sfruttare l'accelerazione delle GPU, e che utilizzano l'attention, per tener conto della rilevanza di altre parti dell'input: per esempio, nella traduzione automatica, il testo originale viene trasformato nella sua traduzione in un'altra lingua, tenendo conto del senso e della corrispondenza con le parole nell'originale.

La traduzione automatica, un caso di successo

La tecnica dei Transformer applicata alla traduzione automatica è stata uno dei più clamorosi successi della disciplina, che ha portato a surclassare in pochi mesi per qualità ed efficienza i precedenti sistemi di traduzione automatica che avevano richiesto anni di sviluppo e messa a punto.

I Transformer hanno poi sbaragliato tutte le altre tecniche usate in precedenza nel campo del NLP in ogni tipo di compito: traduzione, classificazione, riassunto, risposte a domande, analisi di opinioni, inferenza linguistica, ecc.

Basta scorrere la classifica dei sistemi a confronto [su SuperGlue](#), una **raccolta di benchmark di analisi linguistica**, per notare non solo che i migliori fanno tutti uso di Transformer, ma che molti già superano in accuratezza le capacità umane.

I Transformer possono essere adattati a nuovi compiti in modo relativamente semplice con la tecnica del fine-tuning. Si parte da un modello pre-allenato su un'ampia raccolta di testi e gli si

fornisce una raccolta relativamente piccola di esempi del nuovo compito da svolgere e il modello si adatta rapidamente.

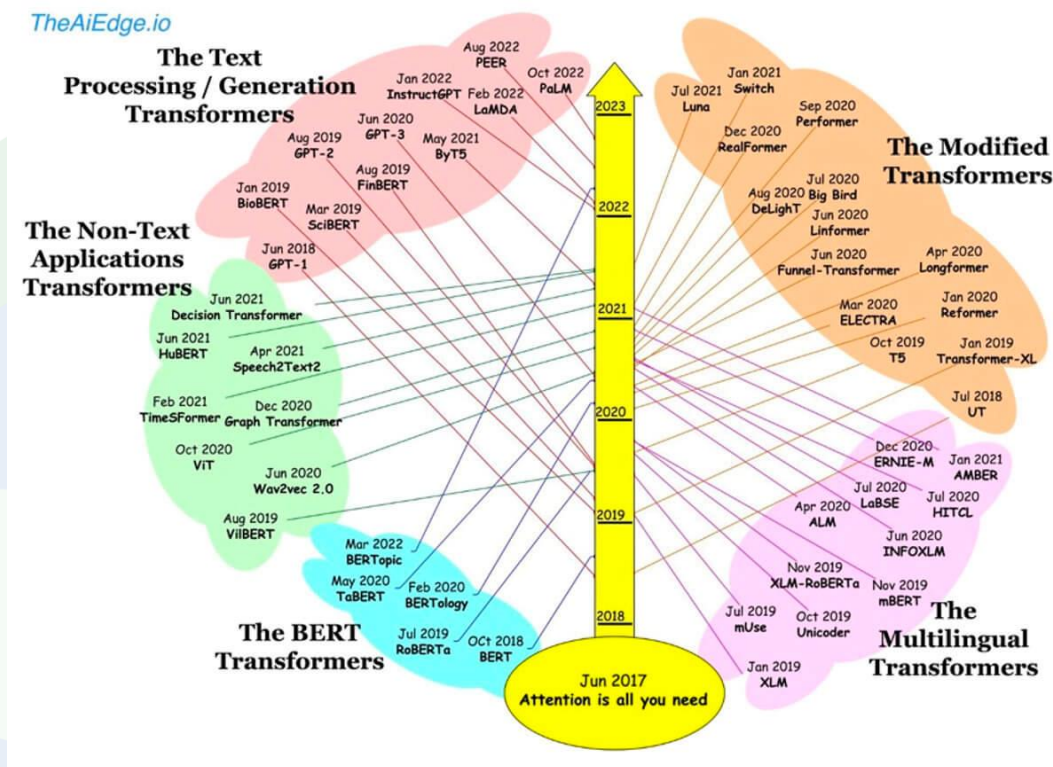
È un progresso notevole perché **accelera lo sviluppo di nuove potenti applicazioni**, sfruttando l'enorme conoscenza linguistica contenuta in un unico Transformer generico, e specializzandolo verso un nuovo compito.

Tutto ciò porta con sé qualità, limiti e – potenzialmente – anche danni.

Il Bello

Transformers History Timeline

TheAiEdge.io



I Transformer fanno parte della Generative AI, sistemi di Intelligenza artificiale **capaci di generare risposte in modo creativo**, producendo risultati che sorprendono per la loro qualità che sembrerebbe tipica della mente umana: testi, immagini, musica e video possono venire generati a partire da frasi che le descrivono.

Per esempio, un testo può essere tradotto a partire dall'originale in un'altra lingua, una figura può essere ottenuta da una descrizione della sua composizione e una musica dal testo di una canzone.

La tecnologia dei Transformer è **applicabile a modalità diverse**, dalla voce ai testi, dalle immagini ai video. Per questo in futuro verranno sempre più sviluppati modelli multimodali, in grado di interagire accettando input sensoriali di tipi diversi e producendo loro combinazioni, rendendo sempre più naturale l'interazione con loro.

I Llm mostrano **risultati impressionanti per una serie di attività di elaborazione** di testi come la risposta alle domande (QA), la generazione di codice (o altri linguaggi formali/assistenza editoriale) e la generazione di storie (fittizie).

Esibiscono capacità che sorprendono gli stessi ricercatori, al punto che sono diventati oggetti di studio per capire quali siano le loro capacità: un settore di studio chiamato BERTology.

Tale studio si esegue stimolando i modelli con delle sonde (probe), per verificare **se sanno svolgere compiti che richiedono capacità per le quali non sono stati allenati**.

I Llm sembrano mostrare capacità emergenti (Wei & al., 2022), ossia **che appaiono solo quando si accresce notevolmente la loro dimensione**, in termini sia di dati di apprendimento che di numero di parametri di cui si compongono. Capacità che non manifestano i modelli di analogia architettura ma di dimensioni più piccole.

Per esempio, **modelli di dimensioni elevate cominciano a esibire capacità di ragionamento** di tipo Chain of Thought, come nella figura accanto, in cui risolve un problema che richiede un ragionamento matematico, seguendo la traccia indicata nella prima domanda/risposta sulle palle da tennis.

Questa sorprendente potenzialità dà ulteriore stimolo a una corsa verso la costruzione di modelli sempre più grandi.

Chain of Thought Prompting

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✓

I limiti

I Llm costruiscono risposte a partire dalle informazioni linguistiche che hanno accumulato nei loro parametri, **non estraggono direttamente la risposta dalle fonti**. Perciò sono utilizzabili per compiti in cui questo modo di operare sia efficace, quali:

- traduzione automatica
- riassunto di un testo
- sintesi di una raccolta di testi
- comporre bozze (di articoli, mail, eccetera)
- trasformare sequenze di un tipo in un altro (testo in immagine, voce in testo, eccetera)

Per questi compiti possono essere di valido aiuto, mentre se si cerca di usarli per ottenere informazioni fattuali **possono cadere in allucinazioni** (hallucinations), introducendo nella risposta informazioni plausibili ma non proprio corrette.

Questo problema viene comunque affrontato con tecniche che guidano un modello a produrre frasi che contengono informazioni precise e corrette ricavate da fonti sicure, ad esempio con la tecnica del prompting, [adottata nei sistemi di data-to-text](#).

ChatGPT è stato allenato a non prendere posizioni su argomenti controversi, e se la cava relativamente bene se interrogato su questioni su cui esiste un'opinione prevalente. Per esempio, **chiedendogli se i vaccini possono causare autismo**, risponde che la scienza è unanime nel negare una correlazione e riporta che gli studi che ne sostenevano l'esistenza sono stati smentiti.

Questo è stato [confermato da esperimenti](#) su vari benchmark, dove ChatGPT risponde correttamente a domande trivia, su fatti che ricorrono frequentemente su internet.

Invece la comprensione del testo necessaria per rispondere a una domanda complessa da un singolo esempio di testo è ancora insufficiente.

D'altra parte, questo significa che **ChatGPT è influenzato dalle opinioni prevalenti** o più diffuse, e quindi non va usato per farsi opinioni o suffragare opinioni preconcepite.

Occorre sempre esercitare il proprio spirito critico e considerare le sue risposte per quello che sono, un'estrapolazione dai testi su cui è stato allenato. **Su molti argomenti non esiste una verità univoca** e non si può certo pensare di trovarla tramite ChatGPT. Una delle stesse fonti principali su cui ChatGPT è allenato è Wikipedia: ma le informazioni riportate su certi temi nella stessa Wikipedia sono il risultato di litigi tra gli editori che cercano di imporre il proprio punto di vista.

Magari in futuro verranno prodotti chatbot che incarnano modi di pensare diversi e gli utenti potranno scegliere a quale di questi aderire per formare le proprie opinioni. Questo però richiederebbe che la capacità di costruire Llm diventasse più accessibile, come diremo più avanti.

I Llm non hanno capacità astratte **quali quella di conteggiare, di fare calcoli, di effettuare ragionamenti logici o di pianificare in più passi**. Per esempio, non sempre sanno calcolare quanto è lunga una parola o disegnare un'immagine con mani a cinque dita o una bocca sorridente con il numero giusto di denti.

Con appositi studi si è osservato che gli attuali Llm da una parte esibiscono davvero competenze linguistiche formali come la conoscenza lessicale e grammaticale, illustrate nella figura accanto, ma dall'altra **sono privi di competenze funzionali** richieste per svolgere calcoli matematici o ragionamento logico. (Mahowald, et al., 2023).

Non stupisce quindi che i professori Gary Marcus ed Ernest Davis [abbiano potuto mostrare](#) come GPT-3 non sia in grado di generare storie coerenti a partire da situazioni che richiederebbero qualche tipo di ragionamento, in base a conoscenze di senso comune.

ChatGPT è stato allenato a gestire dialoghi e quindi a tenere traccia dell'intera conversazione, rispondendo a tono, a volte scusandosi gentilmente se gli si segnala un errore e fornendo una nuova risposta per correggersi. Questo **fa sembrare che ChatGPT apprenda attraverso i dialoghi**: in realtà ciò di cui tiene conto è limitato alla conversazione in corso e non si estende a quelle future.

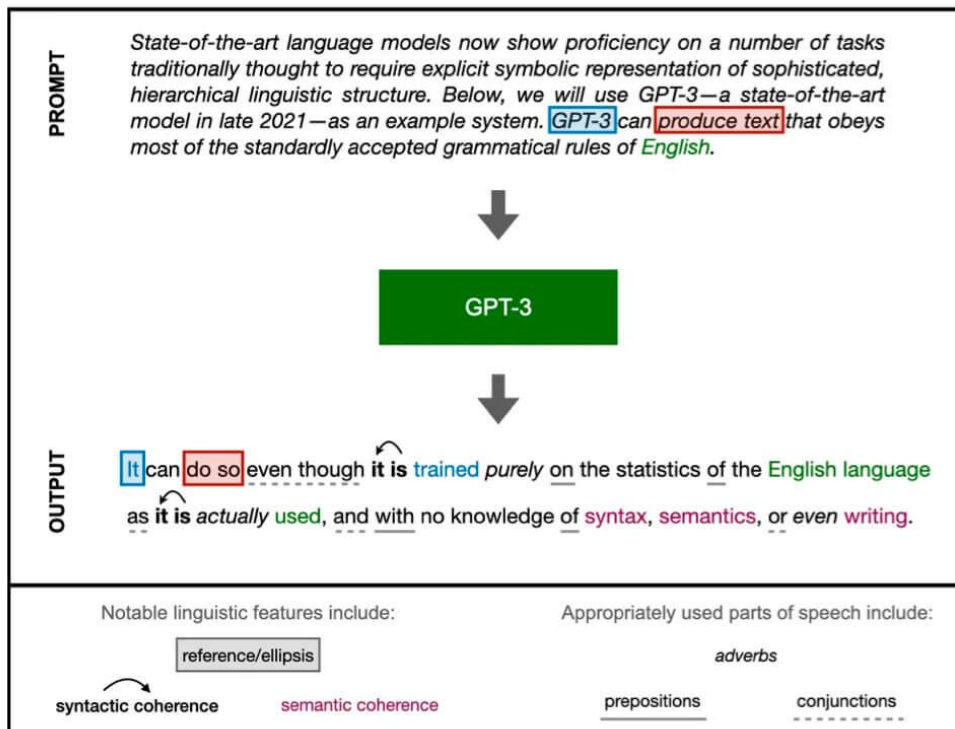
OpenAI sollecita gli utilizzatori a inviare loro feedback sulle risposte, al fine di migliorare il modello, ma ciò avviene con l'aggiunta di nuovi esempi alla raccolta usata per il passo di Reinforcement learning, che richiede settimane o mesi di allenamento e viene fatto quindi solo di tanto in tanto.

ChatGPT ha sollevato perplessità su possibili effetti che il suo utilizzo [potrebbe avere sulla scuola](#), con studenti che si fanno produrre risposte o saggi da ChatGPT esimendosi dallo studio, ma ha sollevato dubbi anche sul mondo dell'informazione e sulla possibilità che possa sostituire i giornalisti nella stesura di notizie.

Altri sistemi [come Dall-E 2](#) potrebbero avere impatti nel mondo creativo, sostituendo gli illustratori con strumenti che generano automaticamente immagini o produrre musica e video. Di recente è stata [minacciata una causa](#) contro l'azienda che produce StableDiffusion, sostenendo che utilizza immagini di apprendimento [ottenute in violazione del copyright](#).

Più grande è il sistema Learning machine, più difficile diventa per l'uomo e per tecniche algoritmiche distinguere le notizie scritte da una macchina dagli articoli scritti da esseri umani e questo apre un dibattito che crea due fronti: da una parte chi opta per bandire l'uso di modelli di AI generative e, dall'altra parte, chi suggerisce metodi di controllo, per esempio con [tecniche di watermarking](#).

I rischi potenziali



È ben noto che le applicazioni di AI generativa come i chatbot a volte possono essere difficili da controllare e si può finire con un programma che vomita commenti razzisti o sessisti. OpenAI ha affrontato questo **problema identificando contenuti tossici** o semplicemente politici generati e fornendo un avvertimento agli utenti.

Di fatto oggi gli unici che possono permettersi le **enormi risorse di calcolo necessarie per allenare un Llm** sono le grandi aziende tecnologiche. E il loro ulteriore sviluppo e diffusione richiede investimenti massicci, come dimostrano i 10 miliardi di dollari che Microsoft ha annunciato di [volere investire in OpenAI](#) e nell'integrazione di ChatGPT con il suo motore di ricerca Bing.

In questo settore stiamo per assistere a una guerra tra titani, per conquistare spazi in un nuovo settore applicativo: Microsoft con [il sistema Prometheus](#) contro [Google con Bard](#).

Il passaggio da un chatbot a un sistema integrato di dialogo e ricerca è tutt'altro che banale, **come emerge dai primi passi falsi di entrambi i sistemi** e richiederà un ridisegno sostanziale dell'architettura del sistema integrato.

Sarà una battaglia cruciale con **effetti dirompenti anche sull'ecosistema digitale del web**: infatti finora i motori di ricerca guadagnavano sulla pubblicità che attraevano sfruttando l'interesse per i contenuti che altri introducevano nei loro siti web. Questi ultimi venivano a loro volta remunerati con un aumento di traffico e una quota di entrate pubblicitarie.

Coni chatbot che producono direttamente le risposte senza fare riferimento alle fonti, si spezza questo cordone ombelicale che alimenta i produttori di contenuti. Gli **effetti di questo cambiamento di paradigma sono del tutto imprevedibili**.

Ci sono due strade possibili per rendere accessibile e democratizzare la tecnologia dei Llm: progetti dal basso che aggregano una comunità di ricercatori come BLOOM (Le Scao & al., 2022) o costruire e rendere disponibili ai ricercatori infrastrutture pubbliche dotate di risorse di calcolo adeguate, come chiedono a gran voce i ricercatori stessi tanto [negli Stati Uniti](#) quanto [in Europa](#).

Conclusioni

I progressi rapidi della tecnologia fanno immaginare ulteriori miglioramenti, sia incrementali sia dovuti ad ulteriori eventuali breakthrough, visto il sempre maggior interesse che la tecnologia sta suscitando nei ricercatori e negli investitori. **Limitarsi quindi alle critiche per gli attuali limiti della tecnologia non tiene conto** che ChatGPT non sarà l'ultimo dei modelli e se si guardasse alla velocità dei progressi recenti, potremmo presto stupirci per nuovi risultati in arrivo.

Occorre però evitare che il controllo della tecnologia rimanga appannaggio di poche grandi imprese che si possono permettere le enormi risorse necessarie per costruire i modelli più sofisticati.

In un caso o nell'altro, l'evoluzione tecnologica dei modelli di AI porterà a cambiamenti dirompenti nel modo di sviluppare applicazioni, nella concentrazione di potere tecnologico e nella disparità tra i detentori della tecnologia e gli altri, e infine nel mondo delle professioni.

Bibliografia

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., (...) Polosukhin, I. (2017). Attention is all you need. Neurips 2017. Curran.

Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2023, January 23). Dissociating language and thought in large language models: a cognitive perspective. Tratto da ArXiv: <https://arxiv.org/abs/2301.06627>.

Le Scao, T., & al., e. (2022). BLOOM: A 176B-Parameter Open-Access Multilingual Language Model. Tratto da ArXiv: <https://arxiv.org/pdf/2211.05100>.

Wei, J., & al., e. (2022, 12). Emergent Abilities of Large Language Models. Tratto da ArXiv: <https://arxiv.org/pdf/2206.07682>.

Il deepfake può anche aiutare nella lotta al revenge porn

Il revenge porn è un problema grave e pervasivo che affligge la nostra società digitale. I deepfake finiranno col fare precipitare ulteriormente la situazione oppure contribuiranno a disinnescare il potere delle immagini intime diffuse senza consenso?

Di **Marco Viola**, Università degli Studi Roma Tre

La diffusione e lo sviluppo delle tecnologie di deepfake spianerà la strada a una diffusione incontrollata della pornografia non consensuale? Vi è sicuramente ragione di allarmarsi e pensare alle contromisure – cosa che infatti alcuni esperti di giurisprudenza stanno facendo. Ma forse, per quanto controintuitivo, è possibile che la diffusione dei deepfake **finisca invece per depotenziare il fenomeno**.

Per motivare il nostro ragionamento, partiamo da un fatto di cronaca che ha destato scalpore e indignazione: il caso della **maestra del torinese** costretta a dimettersi dall'asilo dove lavorava perché alla direttrice erano giunte foto e video intimi che la ritraevano. Gli attori della vicenda che hanno contribuito a diffondere il materiale privato sono stati condannati, ma è difficile pensare che ciò basterà a risarcire la maestra per i danni economici e morali che ha subito in questa faccenda.

È noto, poi, il caso di **Tiziana Cantone**, indotta al suicidio dalla pressione psicologica conseguente alla diffusione virale dei suoi video intimi. È proprio questo episodio che solleciterà la politica italiana a implementare, nella [Legge 69/2019](#) (Codice Rosso), alcune misure specifiche per punire la diffusione illecita di materiale intimo; non un punto d'arrivo, bensì un primo passo.

Revenge porn

Per definire questo tipo di crimini è diventata di uso comune l'etichetta “revenge porn”. Tuttavia, studiosi e attivisti **concordano nel ritenere l'etichetta inappropriata**: soltanto una parte delle condivisioni segue il copione del fidanzato o ex che si “vendica” di una ragazza che si è rivelata infedele o che lo ha lasciato.

Non di rado, infatti, **la condivisione è dettata da altre motivazioni**, che possono andare da quelli che vengono percepiti come “dispetti” [tra compagni di classe adolescenti](#) agli abusi di potere puri e semplici, spesso perpetrati nell'ambito di comunità online che affondano le radici, e a loro volta fomentano, [misoginia e mascolinità tossica](#).

Oltre a ciò, parlare di “vendetta” suggerisce implicitamente che la vittima **abbia commesso qualcosa per cui merita una punizione**. Meglio quindi parlare di pornografia non consensuale o, ancora meglio, di diffusione non consensuale di immagini intime (categoria che include foto e video).

Il fenomeno visto da vicino

La pervasività del fenomeno va oltre le più fosche previsioni. Secondo il rapporto 2022 [pubblicato dall'osservatorio Permesso Negato](#), il fenomeno riguarda “più di due milioni di italiani vittime, **mentre 14 milioni di italiani hanno guardato in rete** immagini di Pornografia non consensuale”.

Impressionanti anche i risultati di un ampio studio internazionale che ha coinvolto 6109 persone di età compresa tra 16 e 64 anni in Australia, Nuova Zelanda e Regno Unito: ben **una persona su tre dichiara infatti che i propri materiali intimi** sono stati sottratti e / o diffusi senza permesso, e / o ne è stata minacciata la diffusione. E ben una su sei confessa di aver sottratto, condiviso o minacciato di condividere materiale intimo.

La stessa ricerca conferma come il problema **colpisce più duramente le categorie più fragili**: le donne, ma anche disabili, appartenenti a minoranze etniche.

Di fronte a dati così preoccupanti, viene da chiedersi se non possa andare peggio di così. E la risposta di alcuni è Sì, forse può andare peggio e il peggio sta per arrivare.

Deepfake e immagini intime non consensuali

Viviamo in un'epoca in cui gli insegnanti di mezzo mondo si stanno domandando come comportarsi di fronte [a compiti svolti con ChatGPT](#) e in cui gli artisti temono la concorrenza degli algoritmi.

Un punto di svolta è stato rappresentato dallo sviluppo delle [Reti generative avversarie](#) (Gan), grazie alle quali un sito come [thispersondoesnotexist.com](#) che può produrre in pochi millisecondi immagini (apparentemente) fotografiche di volti umani difficilmente distinguibili da fotografie reali.

Ovviamente, queste tecnologie hanno investito anche la produzione di contenuti pornografici.

Nel giugno 2019 apparve in rete l'app [DeepNude](#). Sfruttando anch'essa una Gan, l'applicazione **permetteva di “svestire virtualmente” le donne** – ovvero, di trasformare una fotografia di una persona (poco) vestita in un'immagine di donna nuda. Le feroci polemiche fecero sì che dopo solo tre giorni la app venne rimossa dalla rete. Ma il danno ormai era fatto, poiché diversi utenti avevano copiato il codice, che tuttora circola in rete.

Una forma di violenza

Nelle mani di chi traffica in immagini non consensuali queste tecnologie diventano “la nuova frontiera della violenza online”^[1]. Chi condivide immagini e video intime senza il consenso delle persone rappresentate non dovrà infatti nemmeno più fare lo sforzo di procurarsi queste immagini: tutto ciò che occorre sarà un software di deepfake e qualche foto o video (non intimi) delle vittime, spesso di facile reperibilità online.

Ecco che **gli appelli alla “castità virtuale” diventano particolarmente inefficaci**, oltre a rischiare di promuovere nuovamente il victim blaming.

L'unico modo per proteggersi diventerebbe dunque quello di far sparire la propria immagine della rete: niente fotografie su Instagram né sul profilo LinkedIn (ma così chi lo cliccherà più?). Non solo: dovrete convincere i vostri amici a rimuovere quella foto di gruppo delle vacanze dove c'eravate voi... e via dicendo. Più facile a dirsi che a farsi.

Una prospettiva (prudentemente) ottimista

Stando a quanto discusso sinora, dinnanzi a noi **si apre uno scenario piuttosto sinistro**, sembra scontato che la crescente diffusione delle tecnologie di deepfake spianerà la strada a un intensificarsi del fenomeno della diffusione di materiali intimi. Ma se **invece, come ci siamo inizialmente chiesti, il deepfake finisse per depotenziare il fenomeno?**

È quanto sostiene chi scrive e Cristina Voto, ricercatrice e docente di comunicazione visiva presso l'Università di Torino, in un articolo recentemente apparso [sulla rivista Synthese](#).

Suggestionati proprio dal caso di cronaca della maestra d'asilo discusso in apertura, durante gli anni in cui abbiamo collaborato nel progetto europeo [FACETS](#) abbiamo sostanzialmente l'intuizione che, laddove fino ad ora il nostro atteggiamento spontaneo nei confronti di immagini e video con le parvenze di fotografie o riprese è quello di prendere per vero ciò che rappresentano, in uno scenario dove un'immagine o un video generati artificialmente siano indistinguibili da fotografie o video reali e altrettanto (se non addirittura maggiormente) diffusi, **la nostra attitudine sarebbe invece più scettica**.

L'effetto contrario

Qualcosa del genere è stato già osservato da alcuni filosofi del mondo analitico. Per esempio, il professore di filosofia e informatica alla Northeastern University (Usa) Don Fallis ha notato come, in un mondo dove i video fake diventano la norma, la gente **tenderà a diffidare anche dei video reali**, indebolendo il valore informativo di questi ultimi^[2].

Questo indebolimento, che pure per il giornalismo potrebbe rivelarsi una cattiva notizia, potrebbe invece essere un'**ottima notizia nel campo della diffusione** non consensuale di contenuti intimi. Vero, chiunque potrebbe facilmente contraffare e diffondere una fotografia o video che ci ritrae nudi e / o in atteggiamenti intimi, ma proprio questa facilità farebbe sì che il materiale così prodotto non risulterebbe troppo credibile.

Anzi, addirittura la facilità di contraffare siffatte immagini potrebbe indurre a non prendere per vere nemmeno le fotografie e video reali!

Due considerazioni

Il nostro argomento poggia su due assunzioni. Primo, ciò che rende le immagini intime attraenti per chi le osserva – e al contempo espone una vulnerabilità di chi è ritratto – non sono solo le proprietà estetiche dell'immagine, quanto **il senso di connessione diretta che essa genera** con la persona rappresentata.

Da questo punto di vista, le immagini fotografiche sono canali di trasmissione piuttosto efficienti – tanto che in un celebre saggio, il filosofo Walton ci dice che quando guardiamo le fotografie di qualcuno, noi **stiamo al contempo letteralmente guardando** quel qualcuno^[3]. Se questo è vero, allora ricevere una fotografia di nudo nel corso di [una sessione di sexting](#) diventa meno eccitante se si scopre che è pesantemente foto ritoccata, o che addirittura ritrae un'altra persona (pensateci: che effetto vi farebbe scoprire che le commoventi poesie d'amore che vi hanno dedicato sono state realizzate da ChatGPT?).

Un mondo (che crediamo essere) pervaso da deepfake è **un mondo in cui la norma diventa dubitare** che le immagini siano tramite genuini di una qualche realtà rappresentata, fino a prova contraria. O, per lo meno, questo è quanto prevede la nostra seconda assunzione. Un corollario di questa assunzione è che, negli anni a venire, per fare sexting non sarà più sufficiente mandare le proprie foto di nudo: bisognerà dotarle di “sigilli di realtà” indipendenti.

Si tratta di assunzioni corrette? Non possiamo affermarlo con certezza poiché, oltre alle intuizioni (non solo le nostre, ma anche quelle di numerosi altri studiosi), a supporto di entrambe le assunzioni non abbiamo trovato altro che evidenze indirette. Per esempio, a corroborare l'intuizione che uno stesso contenuto mediale sortisca un effetto più intenso quando creduto vero che quando creduto falso, vi sarebbe un vecchio filone di studi psicologici degli anni Settanta sul tema della violenza^[4].

Mentre, a confermarci che lo scetticismo nei confronti delle immagini digitali è possibile - e anzi è già realtà - troviamo alcune interviste a giovani canadesi relative alla loro fruizione dei selfies presenti sui social network, che ci rivelano come “quando si guardano i selfie, di solito si presume l'uso di filtri e decifrare l'autenticità è parte integrante di ciò che si ricerca”. (“When looking at selfies, the use of filters was usually presumed, and deciphering authenticity is integral to what drives looking practices”)^[5].

Indubbiamente sarebbe auspicabile trovare evidenze più dirette. In effetti, stiamo già escogitando dei test più diretti per corroborare o rifiutare le assunzioni che sorreggono la nostra ipotesi. Fino ad allora, ci **pare ragionevole trattarla come un'ipotesi** di lavoro che, pure non avendo (ancora) sufficiente evidenza per dettare l'agenda delle pratiche di contrasto al problema della diffusione non consensuale, abbia però sufficiente plausibilità per immaginarsi che tipo di pratiche si potrebbero trarre. A questo sarà dedicata la prossima e ultima sezione.

Che conclusioni trarre?

Se l'assuefazione ai deepfake potrebbe emanciparci dal potere delle immagini tout court, perché non fare come Mitridate – il leggendario re del Ponte, che a furia di assumerlo in dosi piccole e poi viepiù meno piccole finì per rendersi immune al veleno?

Perché cioè non trarre le conseguenze logiche della nostra predizione e **inondare la rete di svariati fake pornografici**, accelerando così la loro inflazione? Rispondere a questa arguta domanda – che ci fu effettivamente sollevata a commento del nostro articolo dal nostro responsabile di progetto, il semiologo Massimo Leone – richiede di allargare l'orizzonte e prendere in considerazione vari corollari.

Al di là della già menzionata prudenza derivante dallo statuto ancora ipotetico della nostra predizione, un elemento da tenere in conto è che, anche quando manifestamente falsi, **i deepfake possono contribuire a plasmare l'immaginario** visivo di una cultura, cementando o ridiscutendo i confini di ciò che è lecito o illecito, di buono o di cattivo gusto, e via dicendo.

Un potere, questo, che certamente condividono con altre forme di rappresentazione grafica ma che verosimilmente manifestano a un grado più alto, perché anche **ciò che sappiamo non essere reale potrebbe continuare a sembrarci reale**, esercitando così un effetto più profondo sul nostro inconscio.

Pertanto, anche in assenza di pretese di spacciarle per fotografie reali (magari al fine di umiliare una vittima o estorcerle del denaro), elaborare un deepfake a contenuto pornografico di una persona

senza il suo consenso è comunque sbagliato (anche ai sensi di numerose legislazioni), perché ci si appropria indebitamente della sua immagine per connotarla in uno scenario pornografico che potrebbe non esserle gradito. **Il danno però non si estende solo alla vittima stessa**, ma può “dilagare”. In particolare, rappresentare le donne in contesti pornografici può promuovere, in modo spesso subcosciente, il cosiddetto male gaze – un’abitudine visiva che tende a fare del corpo delle donne un oggetto^[6].

La buona notizia è che forse, per insuflare la sfiducia nelle immagini necessaria alla ricetta da noi proposta, non c’è bisogno di utilizzare immagini pornografiche: è verosimile che lo stesso effetto sia sortito **coltivando un sano scetticismo generalizzato verso tutte le immagini** fotografiche o videoriprese.

Uno scetticismo che andrebbe verosimilmente coltivato tramite due canali: uno cosciente che nasce dalla consapevolezza delle possibilità tecnologiche e uno precosciente. Il primo può essere promosso attraverso un potenziamento dell’alfabetizzazione digitale che **non può essere confinato al mondo scolastico** ma è da estendere ai cittadini di tutte le età, mentre il canale subcosciente prevede un lavoro costante su quell’insieme di automatismi e abiti interpretativi che governano il nostro quotidiano, tramite un vero e proprio addestramento a non credere.

Si tratterebbe sicuramente di un percorso lungo e che a livello generale comporterebbe tutti i costi dello scetticismo. Per esempio, per i giornalisti diventerebbe assai più impegnativo far credere che una fotografia o un video siano reali. Ma calato nel contesto della diffusione non consensuale di materiale intimo, questa fatica ricadrebbe sulle spalle di chi traffica in immagini o video (reali o artefatte che siano), ponendo così un ostacolo all’odioso fenomeno.

Certo, anche nella migliore delle ipotesi, perché questa sfiducia possa essere coltivata ci vorranno alcuni anni. Si tratta cioè di una prospettiva di lungo periodo. **Nel breve periodo è però essenziale intensificare gli sforzi** per prevenire il fenomeno e tutelare le vittime.

Inoltre, i trafficanti di immagini non consensuali più determinati potrebbero comunque trovare il modo di superare o aggirare l’ostacolo, escogitando stratagemmi e raccogliendo indizi supplementari per dimostrare la realtà delle immagini fotografiche o dei video che spacciano illecitamente. E forse questo già avviene quando alla diffusione delle immagini intime si accompagna [il cosiddetto doxing](#), il “dossieraggio” relativo alla vittima, che **al danno morale delle immagini aggiunge gravi danni per la privacy e la sicurezza personale**.

Varrà dunque la pena, nei prossimi anni, vigilare attentamente su questo e altri espedienti a cui chi commette questi crimini potrebbe ricorrere per compensare la perdita di “effetto realtà” delle foto e dei video.

Bibliografia

[1] Luca Bainotti e Silvia Semenzin, 2021, p. 57.

[2] Fallis 2021

[3] Walton 1984

[4] Per esempio Geen, 1975

[5] Lavrence e Cambre 2020, p. 5.

[6] Ohman 2020; Harris 2021.

Bibliografia

Bainotti, L., & Semenzin, S. (2021). *Donne tutte puttane. Revenge porn e mascolinità egemone*. Durango

Geen, R. G. (1975). The meaning of observed violence: real vs. fictional violence and consequent effects on aggression and emotional arousal. *Journal of Research in Personality*, 9(4), 270–281

Goodfellow, J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, J. (2014). Generative adversarial nets. *Advances in neural information processing systems*. <https://doi.org/10.48550/arXiv.1406.2661>

Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology*, 34(4), 623-643

Harris, K. R. (2021). Video on demand: what deepfakes do and how they harm. *Synthese*, 199(5-6), 13373-13391

Lavrence, C., & Cambre, C. (2020). Do I look like my selfie?": Filters and the digital-forensic gaze. *Social Media + Society* <https://doi.org/10.1177/2056305120955182>

Öhman, C. (2020). Introducing the pervert's dilemma: a contribution to the critique of Deepfake Pornography. *Ethics and Information Technology*, 22(2), 133-140

Semenzin, S., & Bainotti, L. (2020). The use of telegram for non-consensual dissemination of intimate images: gendered affordances and the construction of masculinities. *Social Media + Society*, 6(4), 2056305120984453

Viola, M., & Voto, C. (2023). Designed to abuse? Deepfakes and the non-consensual diffusion of intimate images. *Synthese*, 201(1), 1-20

Walton, K. L. (1984). Transparent pictures: On the nature of photographic realism. *Critical Inquiry*, 11(2), 246–277

Intelligenza artificiale, che ne sarà dell'uomo? Le riflessioni della Chiesa

L'attualità delle riflessioni sull'Intelligenza artificiale non può prescindere dalla visione dell'uomo che scaturisce dall'insegnamento della Chiesa cattolica e dall'approccio che questa ha maturato, lungo gli ultimi decenni rispetto alla tecnologia e alle sue innovazioni più sofisticate

Di **Giovanni Tridente**, docente di Position Papers alla Pontificia Università della Santa Croce

Di fronte agli sviluppi nel campo dell'innovazione, si fa pressante la domanda legata alla **responsabilità** – che poi del resto è sempre personale, individuale e mai tecnica o artificiale – e alla necessità di “governare” gli strumenti tecnologici. Ed è questa la nuova frontiera della discussione sull'intelligenza artificiale, al di là del funzionamento e della composizione strutturale dei molteplici artefatti. Frontiera su cui si è misurata, almeno a partire dagli ultimi decenni, anche la **Chiesa cattolica**.

Le macchine sono in grado di pensare?

Ogni disamina sull'Intelligenza artificiale (AI) parte inevitabilmente da un riferimento pioneristico legato al famoso test di Alan Turing, il matematico inglese che nel 1950 propose un metodo per valutare l'intelligenza di macchine computazionali messe a confronto con l'intelligenza umana.

Il test prevede che un esaminatore ponga domande attraverso una telescrivente a un uomo e a un computer, situati in due stanze chiuse e distinte, senza sapere se le risposte che riceve vengono dall'uno o dall'altro.

Se l'esaminatore non è in grado di identificare, sulla base delle risposte ricevute, quale sia l'uomo e quale sia il computer, le capacità di interazione linguistica del computer dovranno essere considerate come non distinguibili da quelle umane, e il computer dovrà dunque essere considerato **intelligente**.

L'altro padre dell'intelligenza artificiale è il matematico e informatico Marvin Minsky, a cui è attribuita la prima definizione della disciplina.

Sette decenni fa, dunque, i dibattiti sulle interazioni tra macchine ed essere umano tentavano già di rispondere alla domanda se le prime fossero in grado di “pensare”, di avere dunque una “intelligenza” propria, che si sarebbe messa in **competizione** con quella dell'uomo. E l'interrogativo veniva posto soprattutto rispetto all'intelligenza di tipo linguistico.

Le successive riflessioni e gli studi anche nel settore della sociologia hanno riguardato le previsioni su come sarebbero evoluti i processi interattivi uomo-macchina, fino a desiderare di affidare alla

tecnologia evoluta un'ampia gamma di processi diagnostici, valutativi o **decisionali** che la stessa persona deve svolgere.

Non si può dire che una simile previsione non si sia **avverata**, certamente il percorso con cui si è arrivati a questi risultati ha impiegato diversi decenni. Gli scienziati computazionali dei primi anni Ottanta del secolo scorso, infatti, erano interessati a scoprire “come lavora il cervello umano, come memorizza e come richiama i ricordi, come collega le sensazioni, come cataloga le immagini”, tutto ciò finalizzato alla **creazione** di potenti apparati tecnologici che potessero creare dei “sosia della mente di un uomo ed esporli a determinate sfide e rischi, in processi di controllo e di decisione”.

Già in quelle riflessioni era evidentemente presente anche tutto il tema della “coscienza” e la domanda molto concreta se un calcolatore o un robot potesse essere veramente **cosciente**. Gli sforzi degli scienziati sono andati ovviamente nella direzione di ridurre il divario tra le simulazioni di AI e l'intenzionalità soggettiva e cognitiva di ogni soggetto umano, e hanno dovuto prevedere anche successive implicazioni riguardanti ad esempio l'**empatia**, un elemento fondamentale umano e personale, e da sempre tallone d'Achille di ogni evoluzione tecnologica avanzata, che proprio per questo rischia di non diventare mai completa, almeno nel senso in cui si vuole interpretare l'AI.

Roger Carl Schank, capo del laboratorio di Intelligenza artificiale dell'Università di Yale, già nel 1984 affermava senza mezzi termini che il livello di [completa empatia](#) della comprensione “sembra essere del tutto fuori tiro del computer per la semplice ragione che il computer non è una persona”.

Successivamente anche un altro studioso, John Searle, sostenne che le macchine non possono pensare poiché sono incapaci di recepire i significati e sono prive di intenzionalità.

Tutte le ricerche successive hanno cercato di smentire queste considerazioni puntando a quella che è stata poi definita in gergo come [AI di tipo “forte”](#): ci vorrà del tempo ed energie adeguate ma un giorno secondo alcuni sarà possibile **riprodurre** le attività intelligenti che caratterizzano il comportamento umano in tutte le sue svariate forme.

Call for Ethics in AI, firmato documento da rappresentanti di tre religioni

I rappresentanti delle tre religioni abramitiche hanno firmato il 10 gennaio la call di Roma per l'etica dell'intelligenza artificiale, un [documento](#) pubblicato dalla Pontificia Accademia per la Vita e promosso dalla Fondazione RenAIssance nel tentativo di promuovere l'algoretica, cioè uno sviluppo etico dell'intelligenza artificiale.

Mons. Vincenzo Paglia, presidente della Pontificia Accademia per la Vita, che per primo ha promosso l'Appello di Roma nel febbraio 2020, ha partecipato alla cerimonia ufficiale di firma insieme al rabbino capo Eliezer Simha Weisz (membro del Consiglio del Gran Rabbinate di Israele) e allo sceicco Abdallah bin Bayyah (capo del Forum per la Pace di Abu Dhabi e presidente del Consiglio Fatwa degli Emirati Arabi Uniti).

Per raggiungere gli obiettivi etici nell'AI, secondo quanto espresso nel documento, è necessario partire fin dall'inizio dello sviluppo di ogni algoritmo con una visione "algoritmica", cioè con un approccio di etica progettuale.

Progettare e pianificare sistemi di IA di cui ci si possa fidare implica la ricerca di un consenso tra i decisori politici, le agenzie del sistema delle Nazioni Unite e altre organizzazioni intergovernative, i

ricercatori, il mondo accademico e i rappresentanti delle organizzazioni non governative sui principi etici che dovrebbero essere integrati in queste tecnologie.

Per questo motivo, i promotori dell'appello esprimono il desiderio di lavorare insieme, in questo contesto e a livello nazionale e internazionale, per promuovere l'"algor-etica", ovvero l'uso etico dell'IA come definito dai seguenti principi:

- - Trasparenza: in linea di principio, i sistemi di IA devono essere spiegabili;
- - Inclusione: le esigenze di tutti gli esseri umani devono essere prese in considerazione affinché tutti possano trarre beneficio e a tutti gli individui possano essere offerte le migliori condizioni possibili per esprimersi e svilupparsi;
- - Responsabilità: coloro che progettano e impiegano l'uso dell'IA devono procedere con responsabilità e trasparenza;
- - Imparzialità: non creare o agire in base a pregiudizi, salvaguardando così l'equità e la dignità umana;
- - Affidabilità: I sistemi di IA devono essere in grado di funzionare in modo affidabile;
- - Sicurezza e privacy: I sistemi di IA devono funzionare in modo sicuro e rispettare la privacy degli utenti. Questi principi sono elementi fondamentali di una buona innovazione

L'impatto delle tecnologie sull'individuo

Se guardiamo alle discussioni tra studiosi dell'ultimo ventennio del 1900 sui temi dell'IA ci si rende conto che erano **già presenti**, alcune in fase embrionale, altre più sviluppate, le tematiche principali dell'impatto delle tecnologie evolute sull'uomo e sulla società in generale.

Proprio su quest'ultimo aspetto, il sociologo (e politico) Achille Ardigò tematizzava nel 1986 la necessità di riflettere sulle implicazioni di ciò che la sociologia denomina come "controllo sociale", avanzando **tre ipotesi** su questo legame.

Innanzitutto, l'AI sarebbe (stata) "**causa** di controllo sociale", il cui aumento sarebbe l'effetto dei progressi tecnologici, ma anche "**effetto** di accresciute necessità di controllo sociale" generando investimenti pubblici e privati per rispondere a nuove esigenze sia in ambito civile, di mercato oppure militare e «**correlata** (in tutto o in parte) al controllo sociale e viceversa".

Per avvalorare queste ipotesi, lo studioso cita alcuni "fatti", tra i quali i finanziamenti pubblici negli Stati Uniti da parte del **Pentagono** o incentivazioni da parte di agenzie Nato, o ancora investimenti per la diffusione di sistemi di AI nell'ambito dei [controlli fiscali](#), senza tralasciare il settore medico-sanitario e quello della rappresentanza democratica (elezioni, propaganda, eccetera). Tutti "supporti indiziari molto consistenti", come li definisce Ardigò.

Altri temi delle speculazioni di quegli anni, al di là dell'organizzazione strutturale e di funzionamento dei cosiddetti "sistemi esperti", affrontano già il loro utilizzo nel campo della "salute" (sia sociosanitario che di ausili medici strumentali e [di diagnostica](#)) e anche sul piano della "generatività", intesa come capacità di questi sistemi di ristrutturarsi internamente, propensi a produrre nuove "organizzazioni" di sé stessi.

Che ne sarà dell'uomo?

Trent'anni dopo, la domanda non è più se queste applicazioni tecnologiche così in evoluzione saranno veramente in grado di passare all'ambito operativo concreto, ma se le persone non abbiano piuttosto ormai ceduto ad esse il **monopolio** delle proprie esistenze, creando agli uomini non pochi problemi o sfide.

Di fronte alla pervasività di tutti gli strumenti tecnologici che abitano il nostro presente e che ci assistono in tante operazioni quotidiane (dal comunicare con i familiari al raggiungere una destinazione, dall'esprimere un'opinione a presentare la dichiarazione dei redditi) gli interrogativi di fondo che si palesano riguardano il **“che ne sarà dell'uomo”**: sarà (oppure già è) più libero o più schiavo? più umanizzato o più alienato? più socializzato o più solo? comprenderà di più o sarà ancora più smarrito?

Nel 2020 **Stefano Quintarelli**, imprenditore informatico e già deputato alla Camera scriveva: “Senz'altro, sono stati accesi i riflettori su nuove e più sofisticate tematiche, oltre a quelle evidenziate poc'anzi, tra cui la pervasività degli assistenti vocali di ultima generazione (vedi Alexa, Google Assistant & co.), definiti dei veri e propri maggiordomi, i veicoli a guida autonoma, la digital agriculture, le più sofisticate armi letali anch'esse autonome, le conseguenze sul **mondo del lavoro**, dove in ogni caso gli allarmi sulla fine del lavoro vanno ridimensionati, perché più che l'impiego viene a mancare la specializzazione per certi tipi di attività” (Quintarelli, 2020).

Le preoccupazioni della Chiesa

Ma cosa è l'uomo? Ecco una delle domande che rimandano alla notte dei tempi. La tradizione biblica ne accenna una risposta alquanto dettagliata in riferimento anche al Creatore dell'uomo, risposta che è possibile rintracciare nel Salmo 8, tratto dalla più ampia raccolta di 150 preghiere che nella Bibbia ebraica e cristiana vanno sotto il nome di Libro dei Salmi.

In questo testo, in particolare, mentre si accenna alla magnificenza del Creato, l'autore si domanda cosa possa mai essere l'uomo, creatura indubbiamente molto più contenuta dell'estensione del cielo e ancor meno appariscente rispetto alla luna e alle stelle.

Eppure, riflette l'orante, questo “figlio dell'uomo” è stato fatto “poco meno degli angeli”, coronato “di gloria e di onore” e – cosa ancora più eccezionale – dotato di un esclusivo “potere sulle opere” di tutta la creazione, che viene interamente posta “sotto i suoi piedi”.

Insomma, una creatura così **fragile** ed esigua viene dotata di una dignità suprema. Il punto è che questa concessa signoria potrebbe essere interpretata in maniera errata ed egoistica, trasformando l'uomo in un “folle tiranno” più che in un “governatore saggio e intelligente”, come ha commentato in una delle [sue catechesi](#) Papa Giovanni Paolo II – Karol Wojtyła.

Qui si inserisce tutta la riflessione che la Chiesa e il Magistero ecclesiastico hanno fatto lungo i secoli a proposito della custodia dell'armonia e della bellezza di tutta la creazione che Dio ha affidato all'uomo, affinché “ne usi ma non ne abusi, ne faccia emergere i segreti e sviluppare le potenzialità”, come ha ribadito Papa Giovanni Paolo II.

In questa cornice è possibile intravedere anche tutta la “responsabilità” che deriva all'individuo nel cammino di **comprensione** e sviluppo delle tecnologie in generale e dell'Intelligenza artificiale in particolare, considerate senza dubbio anch'esse frutto di un ingegno primigenio che è stato instillato nell'impronta specifica di ogni essere terreno.

I filosofi, più che domandarsi “cosa è l’uomo?” si sono chiesti piuttosto “**chi è l’uomo**”. La risposta a questo interrogativo pone senz’altro le premesse per una giusta comprensione e utilizzo di tutte le potenzialità che gli sono state affidate, preoccupazione che sta a cuore all’insegnamento della Chiesa, come si diceva poc’anzi.

In questa prospettiva si inserisce la concezione “**personalista**” dell’essere umano, che ha cominciato a diffondersi con l’avvento del cristianesimo e a cui ha dedicato molti studi il filosofo Karol Wojtyła, prima ancora di diventare Papa. Evidentemente, lungo il suo Magistero ne ha potuto accentuare ancora di più l’insegnamento.

Per Wojtyła, la persona è “qualcosa di più della natura individualizzata”. La sua pienezza, infatti, ha a che fare piuttosto con “unicità e irripetibilità”. In questo senso, l’uomo è inteso come «soggetto dell’esistenza e dell’azione». Ciò ne identifica anche la piena dignità, configurando l’individuo come “degno di tutto ciò che gli è dovuto per natura”, uno status che non si perde poiché si è preziosi “in sé” e non soltanto per gli altri.

Resta la possibilità, come sottolinea Aparecida Ferrari, professore associato di Etica applicata presso la Pontificia Università della Santa Croce di Roma, di poter perdere, mediante un esercizio cattivo della libertà, la **dignità morale**, che comunque non cancella quella sostanziale.

Quale visione dell’uomo?

Fatte queste premesse, risulta più semplice addentrarsi nel percorso di analisi della visione dell’uomo che scaturisce dall’insegnamento della Chiesa e dell’approccio che questa ha maturato lungo gli ultimi decenni rispetto alla tecnologia e alle sue innovazioni più sofisticate.

Questo “cammino” prevede senz’altro i pronunciamenti degli ultimi Pontefici in tale ambito. Per il loro legame sia con il mondo ecclesiastico che con quello dell’editoria e della comunicazione, va anche individuata una traiettoria comune che emerge dalla lettura di specifiche pubblicazioni e rubriche apparse negli ultimi anni su testate specializzate (per esempio “La Civiltà Cattolica”).

Non sarebbe completo un simile lavoro di analisi senza una raccolta documentale dell’insieme delle iniziative realizzate nei tempi recenti da alcuni Organismi della Santa Sede dedicate espressamente al tema dell’AI, come la Pontificia Accademia per la Vita o le Pontificie Accademie delle Scienze e delle Scienze Sociali.

Una rassegna di questo dinamismo è stata [raccolta nel libro](#) Anima Digitale. La Chiesa alla prova dell’Intelligenza Artificiale. I risultati emersi dimostrano che l’antropologia cristiana è portata a ribadire che l’uomo, oltre ad essere stato creato a immagine e somiglianza di Dio, racchiude non solo la dimensione corporale ma anche quella dello spazio e del tempo, capace di vivere una serie di **esperienze** che lo rendono unico tra gli altri esseri viventi.

C’è poi la questione dell’agire libero con decisioni prese in maniera cosciente e consapevole, e ciò può avvenire solo nelle relazioni con altri simili. A differenza delle macchine, insomma, l’uomo è in grado di mettere anche in discussione i **principi** e i criteri con cui decide, fermo restando che questa dinamica la può applicare nel momento in cui progetta e programma un sistema di AI, pure avocando a sé tale caratteristica cognitiva propria. In tutto ciò si gioca, praticamente, la dignità umana.

Giovanni Paolo II, oggi Santo, con la sua formazione di indole filosofica, ha puntato evidentemente di più sulle questioni antropologiche, invitando a porsi di fronte alle rivoluzionarie scoperte della tecnica “con **vigile** allenamento all’ascetica”, ben soppesando la responsabilità sociale e internazionale rispetto agli stessi progressi. È il Papa che suggerisce inoltre libertà di giudizio e scelta, uno spirito di servizio e la salvaguardia della dignità dell’uomo.

Con **Papa Francesco** si è già nell’epoca in cui le scoperte sono maggiormente evolute, per cui l’inquietudine della Chiesa assume più un carattere sociale, pur ribadendo con insistenza l’utilità di discernere questi cambiamenti “con coscienza morale” e regolamentarne l’impiego.

Attenzioni espresse anche dalla pubblicistica cattolica la quale, nel rivendicare per la Chiesa un ruolo di primo piano nel sensibilizzare su questi temi, ribadisce la costante **centralità** della persona umana in tutto il processo, la salvaguardia della sua dignità e la necessità dell’inclusione.

Bibliografia

Turing, A. (1959). Computing Machinery and Intelligence. *Mind*, 433-460

Horgan, J. (2016, gennaio 30). *In memoria di Marvin Minsky, romantico della scienza*. Tratto il giorno maggio 7, 2022 da le Scienze:
https://www.lescienze.it/news/2016/01/30/news/ritratto_marvin_minsky-2948947/

Ardigò, A. & Mazzoli, G. (1986). *Intelligenza artificiale. Conoscenza e società*. Milano: Franco Angeli

Hofstadter, D.R. & Dennel, D.C. (1985). *L'io della mente*. Milano: Adelphi

Searle, J. R. (1986). *Minds, Brains and Science*. Harvard University Press

Taulli, T. (2019). *Artificial Intelligence Basics: A Non-Technical Introduction*. Monrovia, CA, USA: Apress

Il futuro è arrivato. (2019, dicembre). *Aggiornamenti sociali*, 805-806

Quintarelli, S. (A cura di). (2020). *Intelligenza Artificiale. Cos'è davvero, come funziona, che effetti avrà*. Torino: Bollati Boringhieri

Giovanni Paolo II. (2002, Giugno 26). *Udienza generale*. Tratto da
https://www.vatican.va/content/john-paul-ii/it/audiences/2002/documents/hf_jp-ii_aud_20020626.html

Wojtyla, K. (1982). *Persona e atto*. Città del Vaticano: Libreria Editrice Vaticana

Ferrari, M. A. (2022). *Persona umana*. Roma: Edusc

Tridente, G. (2022). *Anima digitale. La Chiesa alla prova dell'Intelligenza Artificiale*. Todi: TAU

La cancel culture non è un'ideologia: proviamo a comprenderla con la semiotica

La cancel culture viene sovente etichettata come ideologia, ma in realtà si riferisce a discorsi associati a pratiche, call-out e canceling, la cui contestualizzazione trae origine all'interno di comunità online. Ecco come è possibile descriverla, comprenderla e spiegarla con la semiotica

Di **Marco Giacomazzi**, Università di Bologna e **Francesco Barchiesi**, Laureato in semiotica

Nel discorso pubblico si fa spesso riferimento alla [cancel culture](#) come una forma di ideologia politica – una sorta di esasperazione violenta di progressismo – promossa dalle nuove generazioni nei confronti di figure o testi ritenuti problematici indipendentemente dalla loro collocazione temporale, sia che si tratti di personaggi del mondo dello spettacolo a noi contemporanei, o opere d'arte appartenenti a diversi periodi storici.

Ma la cancel culture non è un'ideologia, bensì si riferisce a discorsi associati a pratiche, call-out e canceling, da contestualizzare nell'ambito delle comunità online. Un'ipotesi critica aiuta a comprendere il travisamento di un'espressione. Ecco come è possibile capire la cancel culture con la semiotica.

La disambiguazione di cancel culture

Proviamo a proporre una disambiguazione del significante “cancel culture” attraverso la contestualizzazione di diverse pratiche – specialmente quelle del call-out e del canceling – nate all'interno di comunità online, intorno a specifiche condizioni e avvenimenti di interesse sociale, politico e culturale. In questa maniera, proviamo a spiegare come sia possibile descrivere, comprendere e spiegare il discorso sulla cancel culture, così come le pratiche ad esso associate, in termini semiotici, e nello specifico nei termini di una descrizione dei codici delle subculture digitali che abitano determinati spazi online.

Attraverso una prospettiva transmediale ^[1] – che osserva una circolarità tra i cosiddetti media tradizionali, come i quotidiani di informazione, ed i media digitali –, si potrà far emergere come questi ambienti, pur non costituendosi come luoghi altri rispetto al discorso pubblico o alla vita offline, siano caratterizzati da codici e linguaggi specifici, e come nella migrazione di questi ultimi da un ambiente mediale ad un altro possano avvenire delle traduzioni non sempre fedeli ai contesti di emergenza.

In principio fu il #gamergate

Per proporre una contestualizzazione mediale del fenomeno culturale delle pratiche di canceling - e come si è arrivati a costruire una più generica e confusa cornice di cultura della cancellazione, è necessario partire dal 2015. Fu l'annus horribilis della cosiddetta alt-right, comunità “memetica” di

estrema destra la cui attività di produzione di contenuti online avrebbe poi svolto un ruolo nell'elezione di Donald Trump dell'anno seguente.

Angela Nagle (2015) in *Kill All Normies* mostra i legami tra l'imageboard 4chan e l'espressione di ideologie estremiste come il neonazismo ^[2]. Se le forme più estreme di queste ideologie trovano uno spazio di espressione privilegiato in zone 'periferiche' del web come 4chan, in spazi più mainstream come quello di Twitter si può osservare come si distinguessero già community caratterizzate dall'azione dei cosiddetti troll.

Questi ultimi, attraverso profili più o meno anonimi, contribuiscono alla costruzione identitaria dell'alt-right attraverso episodi di hate speech e cyber-bullismo di gruppo contro quelle che sono riconosciute "minoranze", quindi tramite discorsi anti-femministi e conservatori. È proprio nel 2016 che Twitter inizia a promuovere l'espulsione ^[3] degli account alt-right in quanto violanti le policy della piattaforma.

Momento fondativo rispetto alla nascita di queste comunità online è l'episodio del #gamergate. La campagna di demonizzazione indirizzata verso alcune personalità femminili del mondo videoludico nordamericano nasce nell'estate del 2014, dapprima su forum come Reddit ed il succitato 4chan. Da qui si svilupperanno, poi su Twitter, thread radunati sotto l'omonimo hashtag.

Doxxing e shitstorm

Per quanto possa sembrare un episodio che interessa soltanto una nicchia digitale, per certi versi, con la stagione del gamergate si crea un precedente storico importante per la codificazione di pratiche violente online fino ad allora poco grammaticalizzate. Come per esempio il doxxing ovvero la diffusione anonima di informazioni sensibili (dati anagrafici, indirizzo di residenza, mail privata di una personalità) con finalità intimidatorie.

Analogamente, si parla di shitstorm per descrivere la partecipazione in massa – spesso organizzata – ad una discussione o ad un post di qualcuno con insulti, o comunque con finalità derisorie.

Grammaticalizzazione di pratiche violente online

Con "grammaticalizzazione" si intende la progressiva associazione di una pratica culturale ad un codice semiotico, che ne consente una riconoscibilità e una replicabilità all'interno della semiosfera ^[4]. Secondo il semiotico lituano infatti la cultura si organizza attraverso delle sfere di produzione testuale, caratterizzate dalla presenza di diversi linguaggi, i cui significati vengono poi compresi dalle sfere adiacenti attraverso pratiche di traduzione che ne consentono una lettura, interpretazione, e quindi nuova produzione.

Al centro di queste semiosfere si trovano testi fortemente grammaticalizzati - ossia strutturati in maniera particolarmente prescrittiva dai codici che vi soggiacciono - che si costituiscono così come modelli per nuove produzioni di testi.

Pratiche di intimidazione

Questo tipo di pratiche di intimidazione online assomigliano a episodi - con una precisa tradizione storica - organizzati di attacco verso figure pubbliche su base politica.

Attorno a questi episodi - di natura discriminatoria spesso **misogina, razzista o omotransfobica** - si avverte all'interno delle comunità colpite la necessità di organizzare forme di resistenza.

Emergono le politiche identitarie

Dal momento in cui **Twitter stesso non riesce a svolgere un adeguato lavoro di moderazione dei contenuti** che ospita, **alcuni utenti si organizzano dal basso per appaltare il mancato lavoro di policing**, ovvero per provare a **neutralizzare i comportamenti violenti e rendere lo spazio virtuale più abitabile per quelle identità collettive che più spesso vengono prese di mira**.

Ci si ritrova, in questo caso, all'interno del quadro dell'**identity politics** (Heyes, C. 2020). Seguendo la definizione della [Stanford Encyclopedia of Philosophy](#), la **politica identitaria** significa un'ampia gamma di attività politiche e fonda la teorizzazione nelle esperienze condivise di ingiustizia subita da membri di certi gruppi sociali. Invece di organizzarsi esclusivamente intorno a sistemi di credenze, manifesti programmatici o appartenenza a partiti, le formazioni di identità politiche tipicamente aiutano a rendere sicura la libertà politica di specifici sostenitori marginalizzati nel loro contesto più ampio ^[5].

La stagione del #MeToo

Queste comunità online svolgono un lavoro di policing agito sulle piattaforme. Lo spiega la studiosa dei media **Lisa Nakamura**, esplorando la call out culture come venture community management ^[6].

Questo tipo di pratiche di moderazione della comunità dal basso prende il nome da “to call out” (che a livello letterario possiamo tradurre come “fare appello”, segnalare). La pratica di fare appello si riferisce al concetto di accountability, come rivendicazione di un principio di imputabilità di fronte a determinati post o commenti che vengono ritenuti discriminatori, segnalandoli a un tempo alla piattaforma e alla comunità di riferimento.

Nakamura, dalla sua prospettiva intersezionale, interpreta il call-out come un lavoro di cura non retribuito, svolto da determinati attori sociali a profitto (anche economico) della piattaforma. Un lavoro che definisce indesiderato (“unwanted”, ivi), a dispetto di chi considera la pratica del call-out – che molto spesso consiste effettivamente sulla moderazione linguistica – una pedanteria, se non la manifestazione di un atteggiamento censorio. Così si spiegano infatti le invocazioni della libertà di espressione di cui la direzione trumpiana e gran parte del GOP si sono fatte da subito.

A distanza di qualche anno, Duchi (2019) [osserva](#) una progressiva popolarizzazione di questa pratica di segnalazione che a Nakamura appariva, mentre scriveva, per lo più disincentivata. In questo senso, è significativo sottolineare lo **scoppio, nel 2017, della stagione del #MeToo**. La denuncia pubblica su Twitter. Le vittime si fanno vicendevolmente coraggio in una catena di prese di parola (letteralmente: “anch'io” o “anche a me”), del sistema di abusi sessuali perpetrato dagli uomini di potere dello star system hollywoodiano - a partire dal **produttore Harvey Weinstein, passando per altri casi che hanno raggiunto una portata di scala globale, come quello dell'attore Kevin Spacey**.

Dalla dimensione social al mainstream

Il #MeToo ha rappresentato uno scandalo internazionale. Si può osservare come le pratiche pubbliche del call-out escano in questo caso, per la prima volta, dalla dimensione mediale specifica delle piattaforme social e facciano breccia nel discorso pubblico mainstream e globale.

Quando si parla di #MeToo, ci si sta infatti riferendo ad un movimento che nasce come hashtag su Twitter. Prende poi progressivamente spazio in tutti gli ambiti della cultura. Ma anche nel lavoro e nelle istituzioni. È un'occasione in cui diventano tangibili presso la società civile le potenzialità non solo di fuoriuscita e popolarizzazione delle pratiche specifiche agli ambienti digitali, ma anche di una loro istituzionalizzazione.

Così, irrobustita dalla stagione del #MeToo, si viene così a costituire una più ampia “call-out culture” come oggetto discorsivo che può essere usato tanto dagli attivisti progressisti quanto dai detrattori conservatori.

Se si legge questa pratica online all'interno della cornice di senso dell'economia della reputazione dei social media, è facile comprenderne il successo. I singoli utenti potrebbero essere più interessati a beneficiare di un ritorno di immagine a partire dall'occasione di manifestare il proprio valore morale, affermando al contempo la propria appartenenza ad un'identità collettiva - il cosiddetto virtue signalling - che a spendersi effettivamente per una causa sociale.

Questa nozione di economia della reputazione permette, tra le altre cose, di spiegare la migrazione dei codici delle comunità online ai reparti di diversity management in **ambito aziendale**.

Perdere follower su Instagram o Twitter significa un'immediata perdita di reputation, attenzione e dunque fatturato [\[7\]](#).

Slacktivism

Per questo tali pratiche di attivismo online sono spesso tacciate di attivismo performativo, o **slacktivism**, una forma di attivismo che non ha alcun effetto trasformativo, se non un ritorno di immagine, moneta principale di una dimensione di simulacri, sulla realtà e non richiede un effettivo lavoro da parte di chi lo promuove.

Tuttavia, se si adotta una prospettiva semiotica non ci si interroga sulla sfera delle intenzioni che motivano i processi di produzione testuale, e diventa estremamente meno interessante cercare di determinare quanto l'adesione ad una determinata causa sia un gesto privato; sarà più interessante osservare le implicazioni delle affordances della piattaforma nel conferire a una battaglia per la giustizia sociale una forma specifica.

Il Twitter mob

Al contempo, è necessario segnalare come la pratica del call-out, proprio per com'è costituita nella sua natura mediale, possa a sua volta finire per ribaltarsi in una pratica violenta. La partecipazione numerosa ad una discussione attorno ad un hashtag con post, commenti e retweet finisce per creare **l'impressione di una massa inferocita**. La cosiddetta “**Twitter mob**” si accanisce, in nome della giustizia sociale, contro l'espressione di un soggetto.

In una dimensione di collasso del contesto ^[8], qualsiasi enunciato può a tutti gli effetti essere decontestualizzato, allestendo un tribunale popolare che dà luogo ad una forma di “**online public shaming**” paradossalmente accostabile, nelle ricadute, alle **shitstorm** proprie invece delle comunità alt-right, cui si è fatto riferimento in precedenza.

Dal call-out al canceling

Non a caso, è proprio questo aspetto a costituire l'accusa principale che si muoverà contro quella che viene definita la cancel culture: gli “attivisti che sono andati troppo oltre” ^[9].

Per comprendere meglio il rapporto tra le diverse forme di codificazione culturale delle pratiche - che seguono, secondo la lezione della semiotica, modelli di costruzione polemica delle posizioni - **manca un anello di congiunzione**: l'esposizione di un'ulteriore pratica online che prende il nome di **canceling**. Essa verrà **accostata al lessema del call-out** fino a sostituirvisi, nell'uso denotativo che se n'è fatto all'interno dell'opinione pubblica. Ciò succede nel discorso pubblico italiano. Infatti nel nostro Paese si è difficilmente impiegato il termine “call-out”.

Il canceling in una prospettiva transmediale

Se si osservano le intersezioni tra diversi ambienti mediali, all'interno di un contesto di cultura convergente ^[10], si può notare come **il termine canceling sia migrato dalla dimensione televisiva al contesto specifico dei social media**.

Secondo la reporter Aja Romano, la pratica del canceling si diffonde in seno alle **comunità del Black Twitter durante il 2015**. Lo fa attraverso **meme che si riferiscono ad una puntata di un reality show americano**. In questo contesto, una concorrente, per rispondere ad un episodio di sessismo, si riferisce ad un altro con l'espressione “You are canceled!”.

La cancel culture viene dunque vista come un'estensione della call-out culture ^[11].

Esiste una continuità tra la pratica del canceling agita dalla comunità afroamericana su Twitter, ed una storia di opposizione articolata dai movimenti politici. Si fa riferimento qui ai **civil rights boycotts** degli anni Cinquanta e Sessanta. **All'interno di questo frame, l'attività di boicottaggio viene interpretata come un rifiuto di partecipare**. Come l'unico strumento di opposizione individuale - e collettivamente organizzato - di fronte alla constatazione di un problema sistematico quale è quello razziale.

La **logica della resistenza passiva** - propria dell'esperienza politica degli anni Sessanta - del “non ho potere, se non quello di ignorare” si fa resistenza attiva nella sua collettivizzazione.

Il boicottaggio

Appare però immediata una discontinuità tra le varie pratiche online. Come già accennato, l'obiettivo del call-out sarebbe quello di fare appello, convocare l'imputabilità - o **accountability** - del soggetto di fronte ad i suoi errori. Avrebbe quindi uno **scopo pedagogico**.

Il canceling invece si contraddistingue come una forma di protesta e di scontro. Consisterebbe quindi in un tentativo di boicottaggio. Due pratiche, pur avendo una differente origine e morfologia, finiscono per associarsi, una volta imbracciate da una medesima comunità.

Educazione e conflitto sono infatti due corni di una stessa attività di policing, agita da parte di soggettività collettive, all'interno di uno stesso spazio mediale in questi stessi anni. Sono le produzioni di senso di identità politiche iper-codificate, grammaticalizzate, che esercitano un lavoro di cura o un'attività di protesta - forse due facce di una stessa medaglia, non dissociabili fino in fondo - all'interno di Twitter, nella seconda metà degli anni Dieci.

Il caso JK Rowling

Un esempio che può favorire l'intelligibilità del canceling - nella sua associazione con il call-out, e nella successiva risemantizzazione dell'evento come caso di cancel culture - è **quanto avvenuto alla celebre scrittrice JK Rowling, tra il 2019 ed il 2021.**

L'inventrice di Harry Potter sarebbe stata **“cancellata” per essere entrata in polemica con la comunità LGBTQ+ su Twitter**, in seguito alle svariate occasioni in cui avrebbe manifestato, con i suoi post, delle **opinioni percepite come transfobiche**. O più precisamente, avrebbe esplicitato la sua **collocazione nel cosiddetto femminismo TERF**, la corrente radicale che **esclude le donne trans dalle rivendicazioni proprie del femminismo**, insistendo su una **visione del genere come strettamente dipendente dal sesso assegnato alla nascita**.

Queste posizioni, che vengono considerate come casi di hate speech, dal momento che negano la validità dell'esperienza delle persone appartenenti a una categoria marginalizzata, finiscono inevitabilmente per diventare oggetto di discussione. Si innesca un botta e risposta con alcuni account, che non si conclude in maniera pacifica: **JK Rowling viene segnalata alla comunità online per essere transfobica (il momento del call-out).**

La contestazione, radunata dall'hashtag che diviene trending topic su Twitter, si inasprisce. Fioccano commenti d'odio. Vengono postati video di fan delusi che bruciano i propri libri di Harry Potter. E **si chiede ai famosi attori che hanno interpretato i personaggi della saga di dissociarsi pubblicamente** - sulla piattaforma - dalle dichiarazioni di Rowling. Scatta così il boicottaggio della scrittrice **(il momento del canceling)**.

La lettera dei 153 intellettuali

La polemica compie un salto di popolarità circa un mese dopo, quando viene pubblicato su **Harper's Magazine un manifesto firmato da 153 “intellettuali”**, col titolo "[A Letter on Justice and Open Debate](#)". L'iniziativa è sostenuta da personalità pubbliche del calibro di **Noam Chomsky, Margaret Atwood e la stessa Rowling**. E diventa celebre sui quotidiani internazionali ed italiani come “la lettera contro la cancel culture”.

JK Rowling denuncia la sofferenza subita per questa esperienza di boicottaggio, condita da minacce e insulti; al tempo stesso rivendica la propria libertà di pensiero e di opinione - che, per la comunità che l'ha contestata, continua ad essere la rivendicazione di una libertà di discriminare.

Definizione di cancel culture

Si è finora cercato di **ricontestualizzare l'emersione di queste pratiche digitali, che saranno inglobate dal caotico significante “cancel culture”**. All'interno del dibattito pubblico, soprattutto nel contesto italiano, si è sentito parlare di cancel culture come tentativo di “cancellare il passato”: la storia, la memoria, le tradizioni (Figura 1).

[infografica id="174385" class="attachment-full infoImg infoImg-contain-width infoImg-contain-width-transform hide" alt="Cancel culture" caption="Figura 1"] [infografica id="174385" class="attachment-full infoImg infoImg-contain-width infoImg-contain-width-transform hide" alt="Cancel culture" caption="Figura 1"]

Come si sarà notato però, **l'uso sinonimico delle due espressioni – call-out e canceling, che si è cercato di disambiguare** – non presenta davvero il problema della cancellazione o della censura, quanto la necessità di interagire con determinate espressioni di senso che ora tendono ad associarsi a certe conseguenze.

Come si evince dal caso Rowling, e come nota anche Raffaele Alberto Ventura ^[12], un **comportamento percepito come sessista o razzista ad esempio, sui social, tenderà a non passare più inosservato** – come poteva accadere fino a pochi anni prima, in determinati contesti pubblici. Questo è indice di una sensibilità mutata. Ed il mutamento si rende particolarmente visibile nell'acceleratore semiotico delle piattaforme social.

Accountability, boicottaggio, cancellazione

Per spiegare l'allacciamento semantico della segnalazione (richiesta di accountability) e del boicottaggio, con quello della cancellazione, abbiamo formulato un'ipotesi provocatoria. L'imprevedibilità delle logiche enciclopediche avrebbe questa volta mescolato le carte in gioco.

Il termine “canceling” è sorto all'interno di una dinamica transmediale: “You are canceled!”. L'espressione potrebbe benissimo essere tradotto con “Per me è finita!”. Oppure: non voglio più saperne niente, non voglio più avere niente a che fare con te.

In questo senso “cancellare” significherebbe semplicemente “chiudere con qualcuno o qualcosa”: unfolloware. E quindi sottoporre a deplatforming ^[13], boicottare, fare terra bruciata. Questo suo aspetto semantico favorisce un'associazione dell'espressione alla pratica del call-out, finendo progressivamente per rimpiazzarla e divenire riferimento predominante. Il tutto avviene **con la collaborazione dell'appropriazione egemonica da parte dei media conservatori americani.**

L'abbattimento di statue, simboli neocoloniali

Uno dei casi in cui emerge maggiormente questo slittamento semantico è quello dell'abbattimento di statue. Le statue abbattute presentano **simboli considerati neocoloniali.** L'azione simbolica di **smantellamento** della statua lascia spazio all'**interpretazione di una cultura della cancellazione che vuole, in qualche maniera, cancellare la storia.**

Ma, come si è provato a dimostrare, far ricadere questa **eterogeneità di pratiche sotto una stessa etichetta** rischia di **omologare gesti simbolici con connotazioni molto diverse.**

Inoltre, in questi casi, **la cancellazione della storia appare completamente fuori luogo.** Anche il tentativo di spiegare queste pratiche attraverso i riferimenti alla **damnatio memoriae**, come **pratica di riscrittura sistematica delle fonti**, rientra negli stessi processi di **traduzione “aberrante”.**

Se si considera infatti lo spazio come un ambiente significante in cui l'architettura svolge una funzione simbolica ^[14], si vedrà come l'erezione di monumenti sia il prodotto di scelte che rimandano a quadri memoriali specifici, e non a una presunta storia universale. In quanto tali,

fungono da campo di forza per rivendicazioni politiche. Infatti, **nell’abbattere una statua, piuttosto che voler cancellare la storia, si sta cercando di contestare la scelta politica di selezionare una porzione della storia a simbolo protagonista dello spazio pubblico.** Senza considerare il fatto che, in termini semiotici, cancellare un segno è qualcosa di estremamente problematico. Ed ogni tentativo di cancellazione va letto come una nuova produzione di senso ^[15].

È dunque possibile ora tentare, a ritroso, di ricostruire gli snodi connotativi e traduttivi. Quelli che hanno portato a questa omologazione di diverse pratiche all’interno della stessa, caotica, etichetta. E con essi tracciare l’operazione di risemantizzazione della pratica del canceling da parte dei mainstream media americani conservatori.

L'intervento della destra trumpiana

È infatti in questo ambiente mediale ideologicamente connotato che **il concetto di cancel culture viene esplicitamente armato dalla destra trumpiana.** Lo stesso Trump avrebbe ad esempio asserito, in occasione della convention repubblicana del 2020: “L’obiettivo della cancel culture è far vivere i 'decent Americans' nella paura di essere bruciati, espulsi, svergognati, umiliati e cacciati dalla società che conosciamo”.

Come [riporta](#) Romano, chi vota Repubblicano dice di sapere cosa sia la "cancel culture" in misura doppia rispetto ai Democratici e la paragona a violente rivolte politiche.

Alla del 2020, almeno un terzo degli speaker ha fatto della cancel culture un punto centrale del proprio intervento. Si parla di cancellare la storia, ammutolire i cittadini e violare il libero scambio di idee, pensieri e parole.

Woke mob

Nella primavera del 2021 alla CPAC (Conservative Political Action Conference), intitolata proprio “America Uncanceled”, si stabilizza ufficialmente la contro-narrazione della cancel culture come frame conservatore.

L’osservatrice del New York Times, Kristi Noem, nota come questa sia il nuovo cavallo di battaglia di Trump – assieme al concetto di “woke mob” - dopo quello di “fake news”, tormentone agitato nella campagna elettorale del 2016.

Aja Romano [osserva](#) invece come si tratti di un’operazione egemonica di panico morale, in continuità con le strategie di gestione della discussione politica su temi quali la questione del politicamente corretto già a partire dagli anni Novanta ^[16], nel contesto anglo-americano.

Come ha notato la sociolinguista Deborah Cameron ^[17] nella sua esposizione del concetto di igiene verbale, si tratterebbe di una strategia di ridicolizzazione e svilimento delle proposte da sinistra sulla moderazione del linguaggio e sulla rappresentazione delle diversità, che manifesta le sue radici nei differenti precedenti storici che hanno costituito il frame interpretativo della cosiddetta “loony left”: la sinistra strampalata, che amplifica questioni di natura semio-linguistica allontanandosi dal senso comune e dai “veri problemi della gente”.

Conclusioni

Emergono chiaramente gli effetti dei processi traduttivi intorno ai fenomeni all'interno di contesti mediali ideologicamente polarizzati. Nella ricontestualizzazione proposta emerge una forma quasi sclerotizzata della struttura polemica che Lotman individua nell'organizzazione dei testi di una cultura.

A partire dal gamergate emergono pratiche di resistenza come quella del call-out. Ad essa si affiancano dinamiche di boicottaggio - con precise derivazioni storiche - che vengono tradotte di contesto mediale in contesto mediale, fino ad assumere diverse forme e connotazioni nel discorso pubblico. Una volta acquisita una denominazione e una riconoscibilità, queste pratiche possono quindi essere interpretate da parte di diverse comunità in quanto unico oggetto discorsivo, un significativo vuoto che è campo di forze per diverse strategie identitarie di legittimazione politica in conflitto tra loro.

Si vuole quindi porre l'attenzione sul fatto che un altro dei motivi che sostanziano la formazione discorsiva dell'oggetto cancel culture, quanto meno all'interno del dibattito pubblico italiano, è una ragione profondamente mediale.

Se nel mainstream si è in grado di indicare una serie di pratiche eterogenee e diverse tra loro come appartenenti ad un unico quadro significante, è perché nelle pratiche traduttive si sono persi degli elementi di decodifica che ne consentirebbero un'interpretazione felice ^[18]. Un'interpretazione adeguata ai loro elementi costitutivi, co-testuali e circostanziali.

Lo scenario italiano

Nel dibattito italiano sulla cancel culture infatti non si fa riferimento al gamergate. O alle specificità dei diversi codici subculturali appartenenti al mondo digitale. A questo, si aggiunge un filo conduttore - che ha precise ragioni storiche - tra i media americani e quelli italiani. Questi ultimi, per adeguarsi ai prerequisiti sensazionalistici della click economy, si ritrovano a importare notizie scorrette. O inquadrare in maniera fuorviante, come il caso della notizia del bacio di Biancaneve.

Si tratta quindi di un medesimo problema di traduzione, per dirla in termini semiotici. Sia per quanto riguarda i contenuti dell'informazione, che vengono così decontestualizzati nell'importazione presso la semiosfera italiana, sia per quanto riguarda il modello economico della comunicazione informativa digitalizzata, improntata al sensazionalismo e al clickbaiting.

Il quadro discorsivo della cancel culture viene quindi attivato da iniziative potenzialmente scandalose, che una volta riprese si prestano a fare engagement su social come Facebook (più usato in Italia rispetto a Twitter), e che vengono effettivamente considerate come scandalose perché collocate in una dimensione di conflitto socioculturale.

Quest'ultimo si rivela essere un più profondo conflitto tra codici semiotici: è la reazione indignata ad una proposta di senso (magari esasperata, e gonfiata ad arte nella costruzione della notizia) che non si riesce a decodificare, ad interpretare se non in termini polemici; nel migliore dei casi come un'assurdità, nel peggiore come un apocalittico programma di autodistruzione della cultura occidentale di cui siamo vittime.

La volontà politica

Ogni invenzione all'interno del campo della cultura viene descritta da Lotman come un'esplosione [\[19\]](#), il cui riverbero viaggia a diverse velocità all'interno delle semiosfere. E subisce continui processi traduttivi.

Le identità che animano le pratiche del call-out e del canceling pongono una critica ad una tradizione consolidata e a sua volta grammaticalizzata, riconosciuta come imperialista e sessista. Si tratta di saperi che, rispetto alla tradizione cui si rivolgono, hanno una storia recente. Il campo discorsivo dei gender studies emerge ad esempio appena cinquant'anni fa.

Questi saperi presentano a loro volta un alto grado di grammaticalizzazione, soggetto a nuovi codici. Anche queste dinamiche traduttive rientrano nel modello polemico di organizzazione della cultura, dove le diverse semiosfere - caratterizzate da diversi codici - presentano diversi gradi di permeabilità. Da qui deriva la difficoltà ad assorbire queste proposte di senso. La difficoltà di permettere loro di muovere verso il centro della semiosfera - sfondare, in un certo senso, nel mainstream.

Adottando questa prospettiva, un fraintendimento basato su una discrepanza di codici sembra perfettamente plausibile. Nella stessa dimensione polemica di conflitti e negoziazioni, è dunque altrettanto plausibile considerare come strategica la volontà politica di farlo di proposito per opporvisi. E, di conseguenza, sabotarle. Ma, di nuovo, lo spazio dell'intenzione rimane al di fuori dello sguardo semiotico, e questa può rimanere unicamente un'ipotesi critica.

Bibliografia

Jenkins, H. (2006): *Convergence culture: where old and new media collide*, New York, New York University Press, trad.it *Cultura convergente*, Apogeo editore, Milano, 2014. [↑](#)

Sempre da lì nasce infatti la fantasia di complotto conosciuta come QAnon, che si è svelata al mondo con l'assalto del Campidoglio nel gennaio 2021. [↑](#)

"When Twitter conducted a so-called 'purge' of alt-right accounts in 2016", pag. 215, Rogers 2020. [↑](#)

Lotman J. M., 1975: *La semiosfera, l'asimmetria e il dialogo nelle strutture pensanti*, a cura di Salvestroni e Sedda, La Nave di Teseo, Milano, 2022. [↑](#)

Stanford Encyclopedia of Philosophy. [↑](#)

NAKAMURA, L. (2015): "The Unwanted Labour of Social Media: Women of Colour Call Out Culture As Venture Community Management" in *New Formations*, 86, pp.106-112. [↑](#)

LOVINK, G. (2022): *Stuck on the platform*, Valiz, Amsterdam (pp.147-148). [↑](#)

DAVIS, J.L., JURGENSON, N. (2014): "[Context collapse: theorizing context collusions and collisions](#)", *Information, Communication & Society*, 17:4, 476-485, DOI: 10.1080/1369118X.2014.888458. [↑](#)

“Activists have gone too far” è l’adagio popolare con cui, negli States, si cerca di infantilizzare la sensibilità politica dell’attivismo giovanile sui social. [↑](#)

JENKINS, H. (2006): *Convergence culture: where old and new media collide*, New York, New York University Press, trad.it *Cultura convergente*, Apogeo editore, Milano, 2014. [↑](#)

"Cancel culture can be seen as an extension of call-out culture: the natural escalation from pointing out a problem to calling for the head of the person who caused it". [↑](#)

VENTURA, R.A. (2022): “Dieci tesi sul politicamente corretto Nuovi codici e nuovi conflitti” in AA. VV., *Non si può più dire niente?*, UTET, Torino. [↑](#)

ROGERS, R. (2020): “[Deplatforming: Following extreme Internet celebrities to Telegram and alternative Social media](#)”, *European Journal of Communication*, 2020, Vol. 35(3) 213–229. [↑](#)

Eco 1975, *Trattato di semiotica generale*, Bompiani, Milano; Violi 2014, *Paesaggi della memoria*, Bompiani, Milano. [↑](#)

LORUSSO, A.M. (2021): “Cancellare, tra segni e codici”, *Filosofi(e) semiotiche*, pp. 1-7.. [↑](#)

FRIEDMAN, J. (2018): *PC Worlds. Political Correctness and Rising Elites at the End of Hegemony*, Berghahn Books, New York, trad. it *Politicamente corretto. Il conformismo morale come regime*, Meltemi, Milano, 2018. [↑](#)

CAMERON, D. (1995): *Verbal Hygiene. The politics of language*, Routledge, London-New York. [↑](#)

ECO, U. (1979): *Lector in fabula*, Bompiani, Milano. [↑](#)

LOTMAN, J.M. 8 (1992): *Kultura i vzryv, Gnosis, Moskva*, (trad. it. *La cultura e l’esplosione. Prevedibilità e imprevedibilità*, Feltrinelli, Milano, 1993). [↑](#)

Un'AI etica è responsabilità di tutti: le basi per realizzarla

L'avventura scientifica dell'AI è nata più di 60 anni fa. Oggi, siamo a ChatGPT. Per minimizzare i rischi insiti nella tecnologia dobbiamo capire che ruolo questa ha nella nostra società. Serve uno sforzo di consapevolezza e collaborazione per identificare tutti insieme la visione del futuro che vogliamo

Di Francesca Rossi, IBM AI Ethics Global Leader, IBM Research, T.J. Watson Research Center, New York, USA

L'**intelligenza artificiale** (AI) offre moltissime opportunità per migliorare la nostra vita. Ormai tutte le **aziende** riconoscono la sua capacità di ottimizzare le loro operazioni e di **creare nuovi servizi** che prima non esistevano. L'AI è anche usata per **accelerare le scoperte scientifiche** e capire come combattere il **cambiamento climatico**. **Supporta anche molte nostre attività**, sia online che offline. Ad esempio, viene usata nel riconoscimento di email spam e di transazioni fraudolente delle carte di credito, nei sistemi di guida assistita, nei suggerimenti online di libri, film o contatti sui social e nella interpretazione dei comandi vocali.

La pandemia, e il conseguente lockdown, hanno dato una forte spinta alla digitalizzazione della società e l'AI ha supportato questo processo facilitando i servizi digitali personalizzati e flessibili.

Ma da dove siamo partiti? E come siamo arrivati a [ChatGPT](#)? Ma, soprattutto, quali sono i rischi e le implicazioni etiche dell'intelligenza artificiale?

Breve storia dell'AI: ragionamento logico e apprendimento da dati

L'AI è una disciplina scientifica e tecnologica con lo scopo di costruire macchine in grado di risolvere problemi che richiedono intelligenza umana. **L'avventura scientifica dell'AI è nata più di 60 anni** fa e include due principali aree di ricerca, che identificano due modi complementari per insegnare ad una macchina a risolvere un problema.

Nella prima, spesso chiamata l'AI "simbolica" o "logica", i ricercatori e gli sviluppatori capiscono come definire un problema (tramite simboli che descrivono le sue componenti principali) e come risolverlo, poi codificano questa soluzione in un algoritmo o delle regole, e trasformano questa soluzione in **codice software** per comunicarlo alla macchina, che da quel momento in poi saprà risolvere il problema seguendo le regole codificate. **Questo modo di costruire macchine "intelligenti" è adatto quando il problema è ben definito**. Un esempio tipico è trovare la strada migliore per andare da una città a un'altra, problema risolvibile da un classico algoritmo di AI che segue la logica per trovare la soluzione migliore in modo efficiente.

Un metodo diverso è invece basato sull'**analisi di dati e l'apprendimento automatico** (*machine learning*), in cui l'AI riesce a trovare correlazioni e *pattern* invisibili alle persone perché "nascosti" in grandi quantità di dati. Questo approccio, che spesso usa strutture chiamate "reti neurali", perché traggono ispirazione dalla struttura del cervello umano, funziona molto meglio di quello basato su

regole e logica quando il problema da risolvere è troppo vago o ha troppe possibili istanze per poter essere descritto precisamente. Un esempio è **l'analisi di un'immagine per capire se contiene il volto di una persona**, dal momento che possono apparire in un'immagine visti sotto diverse angolature. Questo rende praticamente impossibile definire un algoritmo che sicuramente (o con grande probabilità) riesca a dare la risposta corretta. Per ovviare a questo problema, si forniscono invece tantissimi esempi di soluzione del problema (cioè di immagini e corrispondenti risposte corrette) e si lascia che la macchina generalizzi da questi esempi per riuscire a dare la risposta corretta in una percentuale molto alta di casi anche su immagini mai viste prima. Oltre all'analisi di immagini, altri esempi in cui queste tecniche funzionano molto bene sono la traduzione automatica da una lingua ad un'altra, la comprensione di comandi vocali e i sistemi di conversazione, come [ChatGPT](#) (di cui parleremo più avanti).

Mentre le tecniche di AI basate su ragionamento logico hanno bisogno di **un ambiente "controllato"** in cui le macchine si trovano a funzionare, e sono state usate con successo in molti ambiti (quali la programmazione ottimale dei turni del personale e l'allocazione dei *gate* negli aeroporti), le tecniche di apprendimento automatico hanno dato alle macchine capacità "percettive", quali la possibilità di interpretare correttamente un testo, un suono, un'immagine o un video e anche di generare questo tipo di contenuti. Questo ha permesso di **ampliare enormemente il campo applicativo dell'AI** negli ultimi anni.

Le innovazioni che hanno portato ai sistemi AI di conversazione

Molti lettori sicuramente avranno usato ChatGPT, facendo domande e ricevendo risposte spesso sorprendentemente eloquenti e corrette, e si staranno chiedendo se siamo davvero entrati nell'era di una forma di intelligenza artificiale simile a quella umana e se dobbiamo esserne felici o spaventati. Per rispondere a queste domande, dobbiamo prima capire come si è arrivati a questi sistemi conversazionali.

ChatGPT (e altri sistemi AI conversazionali) è costruito usando **le tecniche più avanzate di machine learning**. Queste tecniche, che abbiamo già descritto prima, hanno iniziato ad avere una vasta applicazione solo negli anni 2000, quando dati e potenza computazionale hanno cominciato a essere disponibili in grandi quantità.

Dopo alcuni anni, il "*deep learning*" (DL) ha migliorato la loro applicabilità introducendo reti neurali a più livelli (da qui il termine "*deep*") e mostrando che computer creati per applicazioni grafiche (le GPU) potevano anche essere utili per altri problemi, come quelli legati all'**interpretazione di immagini e testo**. Nel 2017 è stato poi introdotto il concetto di "*transformer*", che ha permesso alle reti neurali di gestire più velocemente e con maggiore correttezza grandi quantità di dati. Questo ha migliorato significativamente le applicazioni relative al testo, ad esempio nelle traduzioni o nella generazione di riassunti.

Questa tecnica è alla base di **sistemi come GPT** (*Generative Pre-trained Transformer*) che possono essere allenati su grandi quantità di dati, prendono come input un pezzo di testo e generano ulteriore testo in linea con l'input. ChatGPT usa GPT e altre tecniche per creare internamente **un modello del linguaggio umano** e poter quindi avere un dialogo dove riesce a rispondere, in modo praticamente perfetto dal punto di vista grammaticale, alle domande di una persona in dialogo su qualunque argomento. Ha però ancora **importanti limitazioni e rischi**, come la generazione di risposte non corrette o pericolose (ad esempio contenenti un linguaggio aggressivo, razzista o comunque non appropriato), dovute a un meccanismo di apprendimento che si basa solo sui dati e

non su regole di ragionamento logico che permettono di distinguere vero da falso, o di riconoscere parole e frasi inopportune.

ChatGPT è quindi **il risultato di una lunga serie di momenti chiave nella ricerca in AI** (inclusi *machine learning, deep learning e transformer*) che, durante i decenni, hanno permesso all'AI di potenziare le sue capacità. Non è una grande sorpresa per i ricercatori di AI, ma è la prima volta che un sistema in grado di “conversare” su qualunque tema è disponibile a chiunque abbia accesso alla rete, e questa è una novità che cambia significativamente il rapporto tra società e AI. Infatti, l'AI è già presente in praticamente tutte le nostre attività, ma la sua presenza è spesso non percepita dagli utenti dei vari servizi. ChatGPT, invece, lo possiamo usare e ci possiamo conversare. Questa modalità di interazione scatena anche la nostra tendenza naturale ad antropomorfizzare una macchina, percependo una “intelligenza” simile a quella umana se osserviamo la generazione di risposte simili a quelle di una persona. Ci sembra quindi che l'AI abbia raggiunto il livello dell'intelligenza umana, e questo ci sorprende, ci affascina e ci spaventa.

Questa non è la prima volta che abbiamo una percezione simile. Già nel 1997, quando il **programma Deep Blue di IBM** vinse a scacchi contro Gary Kasparov, molti di noi provarono sorpresa e preoccupazione, perchè una macchina era riuscita a vincere contro il più grande campione mondiale a un gioco che per noi umani richiede un alto livello di intelligenza. Ma Deep Blue poteva contare su computer molto potenti e veloci, in confronto al nostro cervello, e su esperti umani che avevano aiutato i programmatori a codificare le migliori politiche di gioco. La sfida di Deep Blue è stata una pietra miliare per l'AI, ma non ha avuto un impatto negativo né sul gioco degli scacchi né sulle opportunità umane di usare la propria intelligenza.

Tornando a ChatGPT, in realtà ha acquisito **un ottimo modello del linguaggio umano, ma non ha un modello del mondo**, e non è stato addestrato a distinguere tra vero e falso o a supportare valori importanti per noi umani. Questo non significa che non possa essere molto utile in svariate applicazioni, ma vanno individuate con attenzione. Ad esempio, IBM sta mettendo a punto una modalità di uso di tecniche come quelle che supportano ChatGPT che si appoggia a dati aziendali specifici per ogni cliente, con l'intento di sfruttare le avanzate capacità di gestione del linguaggio naturale ma anche di mitigare **i limiti e i rischi di un uso generale e poco controllato**, che non sarebbe adatto ad applicazioni ad alto impatto sul benessere (fisico, psicologico e finanziario) delle persone.

Intelligenza Artificiale, rischi e valori umani

I vantaggi di questa tecnologia sono chiari, ma bisogna riconoscere che porta con sé anche alcuni rischi, dovuti in parte alla tecnologia stessa e in parte al fatto che **provoca una trasformazione molto rapida della nostra vita**.

Per individuarli, bisogna prima decidere che ruolo vogliamo che la tecnologia abbia nella nostra società. A mio avviso, il ruolo dell'AI è di supportare e accelerare la crescita e la consapevolezza intellettuale, culturale e sociale delle persone, unitamente alla protezione e al supporto dei valori umani.

Questo è condiviso in pieno dall'azienda in cui lavoro, che già nel 2017 ha definito tra i suoi principi fondamentali che “l'AI deve amplificare l'intelligenza umana, non rimpiazzarla”. Quindi, bisogna capire cosa vuol dire amplificare l'intelligenza umana, per poi poter capire se alcuni usi dell'AI possono mettere a rischio questo principio.

Per l'AI e per la tecnologia in generale, molti valori umani importanti potrebbero correre il rischio di essere impattati negativamente. Ad esempio:

- **Privacy:** tra i rischi più importanti c'è la gestione dei dati, che spesso includono anche dati personali, da parte dei sistemi di AI. Sono quindi centrali le questioni legate alla *privacy*. Questo è particolarmente acuito con l'AI, dato che le tecniche di *machine learning* hanno bisogno di grandi quantità di dati. In Europa, la legge "General Data Protection Regulation" cerca di mitigare questo rischio.
- **Fairness e Inclusione:** è anche importante la questione dell'equità ("*fairness*") e dell'inclusione: se non si usano dati bilanciati, l'AI, allenata su quei dati, potrebbe prendere decisioni che discriminano tra vari gruppi di persone. Inoltre, il suo uso e la possibile eccessiva digitalizzazione, potrebbe escludere alcune categorie.
- **Libertà e controllo sulle decisioni:** le tecniche di *machine learning* sono poco "spiegabili" (spesso non è chiaro come la macchina arriva alla sua decisione, a partire dai dati in input) e questo impatta sia sul controllo umano delle proprie decisioni sia sulla fiducia tra le persone e l'AI, compromettendone quindi la proficua collaborazione e il suo supporto al miglioramento delle abilità umane.
- **Impatto sulla crescita personale e l'innovatività umana:** la capacità dell'AI di generare contenuti (come testo e immagini) introduce anche rischi legati all'impatto sul sistema educativo, sui lavori creativi, sull'apprendimento consapevole di meccanismi di ragionamento logico. Se gli studenti usano sistemi come ChatGPT per scrivere un tema, potranno imparare a creare nuove idee e quindi a contribuire alle innovazioni del futuro?
- **Verità e informazione:** ChatGPT e sistemi simili a volte generano contenuti non veri, pericolosi o inappropriati. Senza un controllo attento, la condivisione di tali contenuti può creare false credenze e scardinare un sistema di rispetto e collaborazione basato sulla verità dell'informazione, sia a livello locale che globalmente.
- **Sicurezza e wellbeing:** l'AI non deve mettere in pericolo la nostra sicurezza fisica e psicologica, e in generale il nostro *wellbeing*.

Non possiamo permettere a una tecnologia, e al suo uso, di danneggiare valori umani fondamentali. Questo è vero per ogni tecnologia, ma l'AI ha due fattori aggiuntivi da considerare. Primo, l'IA a questo punto sembra avvicinarsi ad un comportamento quasi umano, almeno nella padronanza del linguaggio. Secondo, il termine usato per questa tecnologia, Intelligenza Artificiale, ci spaventa per il timore di potere essere rimpiazzati da altre entità, percepite intelligenti come e più di noi. Questi fattori contribuiscono a mettere a rischio il nostro stesso senso di identità umana.

Ecco allora che le strategie di sviluppo e innovazione tecnologica si mescolano con le domande etiche sull'impatto dell'AI e che molti attori della società si trovano a parlare di etica molto più di quanto non facessero in passato riguardo ad altre tecnologie.

L'etica dell'AI: non solo teoria

L'etica dell'AI si occupa di **identificare i problemi etici dell'uso pervasivo dell'AI** nella nostra società e di fornire possibili soluzioni, sia tecniche che non, a questi problemi. E' un campo di lavoro multi-disciplinare e multi-*stakeholder*, dove esperti di varie discipline (quali AI, economia, sociologia, filosofia, psicologia) e vari attori nella società (come aziende, università, politici, associazioni di cittadini e consumatori) lavorano insieme per generare soluzioni che includono principi, linee guida, strumenti tecnologici, sistemi di *governance*, *standard* internazionali, sistemi

di analisi del rischio, e leggi, in un mosaico di azioni complementari che affrontano i problemi da vari angoli in modo da generare soluzioni globali.

Negli ultimi anni, molte aziende, associazioni, governi e istituzioni hanno pubblicato o stanno lavorando ad articoli scientifici, principi, requisiti, regole, certificati o *standard* e leggi, con l'intento di studiare e affrontare alcuni degli aspetti elencati sopra e legati all'etica dell'AI. Alcuni esempi sono:

- **Le Nazioni Unite:** Il convegno annuale “AI for Good”, organizzato da una delle agenzie ONU, mira a capire come usare l'AI per avvicinarci agli Obiettivi per lo Sviluppo Sostenibile delle Nazioni Unite, facendo incontrare gli esperti di AI con chi conosce sul campo i problemi da affrontare;
- **Il Foro Mondiale Economico**, con i suoi gruppi di lavoro sull'AI che coinvolgono esperti di tutto il mondo in rappresentanza di aziende, governi e accademia;
- **la Commissione Europea**, con il suo gruppo di lavoro multi-disciplinare sull'AI e la sua proposta di legge sull'AI, attualmente in discussione sia in Europa che in America;
- **la Casa Bianca**, con l'“AI Bill of Rights”, che indica i diritti umani che non vanno danneggiati dall'IA o altre tecnologie;
- **Singole aziende** come Google, Microsoft o IBM, che definiscono processi e politiche interne per lo sviluppo e l'uso dell'AI;
- **Convegni scientifici come AIES** (AI, Ethics and Society) che presentano innovazioni tecnologiche e socio-tecnologiche ad una platea di esperti di varie discipline;
- **Organizzazioni come l'OECD** (Organizzazione per lo Sviluppo e la Cooperazione Economica), con i suoi principi sull'AI;
- **Il Vaticano**, con la “Rome Call for AI Ethics”, pubblicata nel 2020 e firmata da aziende come IBM e Microsoft, e recentemente (2023) anche da rappresentanti delle tre religioni abramiche.

Un approccio olistico all'etica dell'AI

Tra le principali aziende che sviluppano AI e la applicano in svariati scenari e settori, l'azienda in cui lavoro ha adottato **un approccio olistico all'etica dell'AI**, in cui vengono affrontati tutti gli aspetti con strumenti di vario tipo: principi, progetti di ricerca, strumenti *open-source*, prodotti e piattaforme, metodologie di *governance* ed educazione aziendale e anche collaborazioni con governi, altre aziende, associazioni e istituzioni.

Come ho accennato, **IBM è stata la prima azienda a pubblicare nel 2017 i propri principi per lo sviluppo dell'AI**, individuando nella *trasparenza* e la *fiducia* due proprietà fondamentali e indicando la volontà di creare AI a supporto dell'intelligenza umana e non per rimpiazzarla. Nel 2017 abbiamo anche definito e pubblicato una politica di gestione responsabile dei dati, che dichiara i dati di proprietà dei nostri clienti.

Questi principi guidano **tutte le attività dell'azienda**, a partire dai laboratori di ricerca, dove vengono studiate e sviluppate nuove soluzioni per creare un'AI “affidabile”, definita attraverso quattro linee principali: *fairness* (cioè non discriminazione), robustezza, spiegabilità e trasparenza. I ricercatori creano soluzioni scientifiche e tecnologiche innovative per individuare e correggere il *bias* in AI, per dare all'AI la capacità di spiegare le proprie decisioni anche quando si usano tecniche di *machine learning*, per minimizzare gli errori di queste tecniche e quindi renderle più

robuste, oltre a metodologie per documentare la creazione dell'AI e quindi dare trasparenza al processo di sviluppo.

Studi e programmi sull'impatto dell'AI nel mondo del lavoro vengono effettuati internamente o in collaborazione con istituzioni accademiche. Inoltre, programmi gratuiti come [SkillsBuild](#), attivo anche in Italia, aiutano i giovani ad avere le nozioni per entrare in un mondo del lavoro che è ormai pervaso dall'AI.

Perché un'azienda possa veramente produrre **AI responsabile e affidabile**, è necessaria una collaborazione e un coordinamento tra tutte le divisioni aziendali. Per questo abbiamo istituito nel 2018 un comitato interno per l'etica dell'AI, che io presiedo insieme al dirigente per la privacy dei dati, che include rappresentanti della ricerca, unità di business, comunicazioni, ufficio legale e chi si occupa della *privacy* dei dati, con lo scopo di favorire la condivisione di informazioni, processi e materiale utile sull'AI e la sua etica tra le varie divisioni e di aiutare le unità di business a discutere e risolvere eventuali dubbi etici su soluzioni e prodotti: se una proposta di prodotto non passa il vaglio del comitato etico dell'AI, il prodotto non viene creato.

Inoltre, **un approccio multi-stakeholder è necessario non solo al proprio interno, ma anche a livello globale**. Per questo IBM partecipa o guida le iniziative globali di maggiore impatto sull'etica dell'AI, quali la Partnership on AI (dal 2016), il gruppo di esperti IA della Commissione Europea (2018-2020) e la *Call for AI Ethics* della Pontificia Accademia per la Vita (2020).

Costruire un futuro ideale per l'AI e per noi: tutti possiamo e dobbiamo contribuire

L'AI ha una presenza e un impatto pervasivo nella nostra vita a livello globale nella società e nel mondo. Quindi sta modellando la nostra vita e lo farà sempre di più. Per evitare di lasciarci trasportare in modo passivo verso un futuro che non supporta i nostri valori, serve uno sforzo di consapevolezza e collaborazione per identificare tutti insieme la visione del futuro che vogliamo e le traiettorie che giudichiamo accettabili per raggiungerlo.

L'etica dell'AI supporta questo sforzo globale attraverso un lavoro multi-stakeholder in cui tutti hanno un ruolo importante: aziende, legislatori, accademici, media, società e anche noi singoli individui e utenti di servizi supportati dall'AI.

In quanto ricercatrice di AI, faccio del mio meglio per capire come migliorare questa tecnologia e mitigare i rischi delle sue applicazioni, sia nei miei progetti di ricerca che nei processi aziendali che l'azienda utilizza. Sento anche la responsabilità di far capire ai non addetti ai lavori che cos'è veramente l'AI e che ruolo può avere nel nostro futuro. **Sono consapevole delle opportunità e dei rischi dell'uso di questa tecnologia e sono ottimista**, perché in pochi anni ho visto grandi passi avanti e un incremento esponenziale delle iniziative, sia nell'AI che in un approccio globale al suo uso responsabile.

Ognuno di noi può contribuire, usando la tecnologia in modo consapevole, responsabile e rispettoso degli altri e dell'ambiente. Non affidiamoci alla tecnologia in modo passivo. Usiamola per capire meglio, per affinare le nostre conoscenze, esplorare nuove aree, confrontarci con altre persone, cambiare idea, raggiungere dei compromessi e collaborare. Rendiamo concreto nella nostra quotidianità il principio di migliorare noi stessi attraverso la tecnologia, non il contrario. Tutti possiamo e dobbiamo contribuire a creare un futuro dove l'AI supporti i nostri valori.

Bibliografia

Intelligenza Artificiale e Machine Learning:

[Artificial Intelligence: A Modern Approach](#), 4th Edition, Stuart Russell, Peter Norvig, Pearson, 2020

Deep Learning, Aaron Courville, Ian Goodfellow, and Yoshua Bengio, 2016.

Pattern Recognition and Machine Learning, Christopher Bishop, Springer, 2006.

AI fairness e spiegabilità:

[Bias in data-driven artificial intelligence systems—An introductory survey](#), Eirini Ntoutsi et al., 2019,

A Survey Of Methods For Explaining Black Box Models, [Riccardo Guidotti](#) et al, 2018, <https://arxiv.org/abs/1802.01933>

Il confine del futuro: Possiamo fidarci dell'Intelligenza Artificiale? Francesca Rossi, Feltrinelli, 2019

Leggi sull'IA:

General Data Protection Regulation (GDPR), 2016

The EU AI Act, 202, <https://artificialintelligenceact.eu/>

[AI Bill of Rights](#), OSTP, White House, 2022

[European Commission High Level Expert Group on Artificial Intelligence](#), 2019,

Principi aziendali sull'etica dell'IA:

[Google AI principles](#)

[Responsible AI, Microsoft](#)

[Principles for Trust and Transparency, IBM, 2017](#)

Convegno annuale sull'etica dell'IA:

[AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society](#)

Iniziative globali sull'etica dell'IA:

[Partnership on AI, 2016](#)

[OECD Principles for Artificial Intelligence](#)

[Rome Call for AI Ethics, 2020](#)

[ITU AI for Good Global Summit](#)

[How global tech companies can champion ethical AI, World Economic Forum, 2020](#)

IA e etica in IBM:

[Deep Blue, IBM, 1997](#)

[Data Responsibility @ IBM,](#)

[Trusted AI for Business, IBM](#)

[AI ethics @ IBM](#)

[IBM SkillBuild](#)



La demenza digitale esiste? Impariamo a distinguere, con la ricerca “single case”

“Demenza digitale” è un libro di Manfred Spitzer, neuropsichiatria e professore di Università a Ulm, che ci dà lo spunto per riflettere su come la funzione, di crescita o di stagnazione, che l’oggetto digitale determina su un soggetto va valutata in base alle sue caratteristiche e alla specificità del momento attraversato

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 ricerca in **psicoanalisi** è sempre ricerca “**single case**”, basata sull’insegnamento che possiamo trarre da ciascun essere umano con la propria singolarità; al contrario, **le ricerche a proposito degli effetti della diffusione del digitale** nel nostro mondo e dell’eventuale deterioramento indotto dal digitale tendono a basarsi sovente su **statistiche che stentano proprio a cogliere la singolarità di ogni posizione soggettiva**.

Non a caso, da questi dati viene eliminato qualunque riferimento alle strutture cliniche, lasciando indistinto quanto avviene a proposito degli schermi in **bambini autistici**, in bambini con **debolezza mentale** o che provengono da famiglie con importanti **carenze educative e formative**, in bambini con una struttura di tipo psicotico oppure ancora in bambini con struttura rigidamente ossessiva.

Danni fisici e psicologici da digitale: pregiudizio o realtà?

La recente normativa del Ministero dell’Istruzione e del Merito volta a **proibire l’utilizzo dello smartphone** a scuola ha suscitato un aspro dibattito, più per i modi imperativi e categorici che per il contenuto in quanto non introduce di fatto delle innovazioni radicali a livello attuativo.

Ribadisce che l’uso dello smartphone è o quantomeno dovrebbe essere vietato a scuola. Peraltro, sul concetto di merito si potrebbe aprire un’ulteriore discussione dalla quale ci asteniamo in questa sede. Tale intervento restrittivo e dai toni rigidi attingibili fra le righe di questo documento del Ministero si basa anche e soprattutto sugli Atti Parlamentari della Settima Commissione del Senato della Repubblica; composta da figure di svariati colori politici, frutto di un lavoro biennale che ha occupato undici sedute, ha attraversato diverse esperienze di governo e i momenti più drammatici della pandemia.

Nel giugno del 2021, con **Andrea Cangini** di Forza Italia come relatore, la suddetta Commissione, dopo aver ascoltato per due anni diversi esperti del settore nell’ambito di un’indagine conoscitiva sull’impatto del digitale sugli studenti soprattutto a proposito dell’apprendimento, prendeva posizione a proposito della **diffusione dei dispositivi digitali** fra i bambini. Sono stati ascoltati Paolo Moderato, professore ordinario di psicologia generale, così come il valido pedagogista Raffaele Mantegazza e altri docenti universitari e noti studiosi di età evolutiva; i video di queste audizioni sono agevolmente rintracciabili sul sito del Senato della Repubblica.

Nel suddetto documento, la Commissione riferiva “**danni fisici: miopia, obesità, disturbi muscoloscheletrici**” includendo fra questi addirittura “ipertensione e diabete”; con altrettanta se non maggior preoccupazione elencava i “danni psicologici: dipendenza, alienazione, depressione, irascibilità, aggressività, insonnia, insoddisfazione, diminuzione dell’empatia”. Fonte maggiore di inquietudine apparivano soprattutto **le conseguenze attribuite a videogiochi e smartphone sul piano intellettuale**: questa Commissione del Senato si riferiva a “progressiva perdita di facoltà mentali essenziali che sommariamente chiamiamo intelligenza: la capacità di concentrazione, la memoria, lo spirito critico, l’adattabilità, la capacità dialettica”. In questi termini “lo smartphone non è più uno strumento ma un’appendice del corpo. E’ per questo che risulta così difficile convincerli a farne a meno, a mettere da parte il telefonino almeno per un po’: per loro, privarsene è doloroso e assurdo quanto subire l’amputazione di un arto”^[1]. Constatiamo le parole forti, la lettura della situazione relativa alla digitalizzazione orientata in modo perentorio verso un atteggiamento critico. Verrebbe da stupirsi e da considerare questi interventi come il frutto di una sorta di **pregiudizio ideologico**. Troviamo infine davvero del tutto enigmatico il modo in cui l’utilizzo smodato dei dispositivi digitali potrebbe in qualche modo determinare **effetti clinici drammatici come ipertensione e diabete**.

Il concetto di demenza digitale

Tuttavia, alcuni colleghi hanno sottolineato i rischi di un deterioramento delle potenzialità umane, soprattutto fra soggetti in età evolutiva, a seguito della sostituzione di attività tradizionalmente svolte senza supporti digitali con attività analoghe effettuate con dispositivi digitali. A questo proposito, le difficoltà di apprendimento e l’eclissarsi di capacità mentali, come quelle relative al **mantenimento dell’attenzione e della concentrazione**, sono correlate a un concetto formulato una dozzina d’anni or sono da clinici della Corea del Sud in maniera forte: **la demenza digitale**. La popolazione giovanile coreana fruisce da tempo di molteplici dispositivi digitali. Una decina di anni or sono, quando tale denominazione ha conosciuto un certo successo, le statistiche riportavano dati come quello di un 12% circa di minorenni di quella nazione che manifestavano una severa dipendenza da oggetti quali tablet, piattaforme per videogiochi e smartphone. Nel dettaglio, si indicava in 160.000 circa il numero di bambini fra i 5 e i 9 anni alle prese innanzitutto con la dipendenza dai videogiochi ma anche più estesamente con una certa Internet addiction. I clinici sudcoreani riscontravano in questi ragazzi aumento di disturbi della memoria, dell’attenzione e della concentrazione tanto quanto un certo appiattimento affettivo – emotivo.

Tale concetto, il concetto di una deprivazione delle facoltà mentali indotto dalla diffusione dei dispositivi tecnologici soprattutto fra i nativi digitali, fornisce il titolo a “Demenza digitale”. Si tratta di un libro abbastanza divulgativo ma corposo e ricco di riferimenti a ricerche svolte in varie parti del mondo, scritto da **Manfred Spitzer**, neuropsichiatria e professore di Università a Ulm, in Germania. Spitzer è stato il primo fra gli studiosi ascoltato dalla suddetta Commissione del Senato. Posizioni analoghe alle sue vengono espresse dal peraltro serissimo collega **Nicola Artico**, docente a contratto all’Università di Pisa e direttore di struttura complessa nella ASL Toscana Nord Ovest. Ho approcciato allora la lettura di questo poderoso volume, pagina per pagina. Non nascondo di essermi trovato però, in una sorta di condizione di **perplexità dinanzi a tre elementi specifici di questo libro**: un primo è di carattere temporale; un secondo quanto al riferimento alle neuroscienze che si basa su una frequente confusione, riscontrabile sia in questo volume sia nella letteratura del settore, fra mente e cervello; l’altro relativo al titolo del libro.

Primo argomento. A livello di tempi, un libro sul digitale scritto in tedesco nel 2012 sarebbe già assolutamente datato, considerata la **rapida evoluzione delle dinamiche e delle abitudini in**

questo ambito; sembra proprio di un'altra era geologica dopo l'accelerazione della digitalizzazione del mondo indotta da lockdown, quarantena, confinamenti e coprifuochi del 2020.

Secondo punto da problematizzare. Per quanto concerne il **titolo del volume**, demente ovvero demente significa letteralmente mancante della mente: **mi risulta difficile credere che un nativo digitale sia demente, mancante della mente**. Amo poco il termine di deficit, comune in una certa clinica neurologica; sin dall'inizio della mia pratica come psicoanalista, ho sempre considerato le varie forme di psicopatologia più come particolarità soggettive tali da insegnarci qualcosa anziché come dei deficit. A questo proposito, per portare un esempio molto affine al campo del digitale, credo che un bambino autistico abbia una propria diversità e una propria singolarità ma senza doverla rubricare nel campo del deficit o della de-menza. Si prendano in conto per questo le testimonianze di famosi soggetti autistici come lo scienziato Daniel Tammet e la docente di zoologia **Temple Grandin**.

Terzo fattore critico è quello della **sovrapposizione fra la dimensione mentale e quella cerebrale**. Mente e cervello hanno ben poco in comune. La mente prescinde in larga misura dal cervello così come il corpo si distingue dall'organismo. Cervello e organismo si situano su un piano organico mentre corpo e mente – peraltro accostabili a livello psicocorporeo – implicano il campo del soggetto. Soggetto umano, soggetto influenzato nella sua formazione dalla società, dalle strutture ambientali, economiche, culturali, linguistiche, soggetto frutto di interazioni che variano e si evolvono secondo una storicità dell'inconscio e della psiche la quale non potrebbe non risentire dei cambiamenti sociali. Fra i cambiamenti sociali che incidono sullo sviluppo mentale, hanno evidentemente un posto preminente quelli inerenti alla digitalizzazione del mondo. Non è tuttavia detto che si possa disgiungere in modo manicheo, dicotomico l'oggetto digitale quale oggetto cattivo o persino persecutorio da altri oggetti come buoni oggetti. Vi sarebbe l'eventualità di idealizzare altri oggetti tradizionali come il libro, il quotidiano, l'album musicale, il pallone, il quadro attribuendo soltanto ad essi una funzione di sviluppo positivo a scapito degli oggetti digitali che sarebbero finanche nocivi. In realtà potrebbero semplicemente fornire un'altra forma allo sviluppo della mente.

Differenze di età

Ascoltando il video dell'audizione di **Raffaele Mantegazza** da parte della suddetta Commissione del Senato, ci troviamo a concordare sui rischi relativi all'utilizzo dei dispositivi digitali nella fase di acquisizione della letto-scrittura: in quella fascia di età e in quel momento critico dell'apprendimento, l'utilizzo di strumenti tradizionali come carta e penna o matita si fa preferire in quanto include un percorso di approfondimento maggiore. Spitzer riporta dati che indicano come un persistente utilizzo del computer nella fascia d'età della scuola dell'infanzia ha un ruolo nel determinare disturbi dell'attenzione, difficoltà a riconoscere le lettere e successivamente dislessia. **Fino all'età dell'asilo nido e della scuola dell'infanzia, l'utilizzo di tablet o smartphone andrebbe dunque assolutamente limitato**. Spitzer si riferisce a studi di neuroimaging condotti con la risonanza magnetica funzionale (RMI) che mostrano come il riconoscimento di lettere imparate per mezzo della scrittura tramite matita porta a una maggiore attività nelle regioni motorie del cervello rispetto a quando per questo si utilizza la tastiera. Del resto, è evidente anche per gli adulti come leggere un concetto e trascriverlo metta al lavoro in forma diversa rispetto a quanto avviene con la funzione del PC denominata "copia e incolla". **Una norma ragionevole è che i figli nei primissimi anni di vita assistano soltanto a brevi video, alla TV o sullo smartphone;** per esempio, meglio far loro seguire soltanto a pochi episodi di cartoni animati rivolti a bambini come Peppa Pig, giusto per citare un personaggio celebre i cui episodi basati sulla rassicurante routine

della vita quotidiana della più tenera infanzia e con personaggi stilizzati secondo una semplicità tipica di quella fascia d'età hanno durata compresa fra i 5 e i 15 minuti.

Vi è accordo sul nesso fra disturbi dell'apprendimento o da iperattività nell'infanzia e l'evoluzione di questi soggetti verso le dipendenze in adolescenza o nella giovinezza. La riscontriamo ogni giorno nella nostra pratica clinica. Del tutto apprezzabile ci sembra anche il parere di diverse organizzazioni di pediatri, in molteplici nazioni, da noi già riportato più volte nelle nostre pubblicazioni, circa il limite di tempo dinanzi agli schermi da dare ai bambini a iniziare dal limite di un'ora fino ai tre anni di età. Del resto, anche quando si svolgevano [Didattica a Distanza](#) e poi **Didattica Digitale Integrata** in occasione di lockdown e zone rosse o arancioni, al primo anno della scuola primaria, era consuetudine quella di svolgere soltanto un paio d'ore di lezione in tale modalità online e neppure quelle alla scuola dell'infanzia onde evitare di far restare per un tempo inadeguato all'età i bimbi dinanzi al video. Non sono così convinto, tuttavia, che risulti condivisibile quanto afferma **Angela Biscaldi** quando scrive che “i nuovi media digitali, costitutivamente, indeboliscono l'attenzione, la concentrazione e l'empatia” né tantomeno che “non siano adatti per l'approfondimento critico”. Del resto, si tratta di una posizione affine a quella esposta e pubblicata negli Atti Parlamentari del Senato della Repubblica sopra citati.

La Commissione in questione ha audito pure **Lamberto Maffei**, il quale ha insegnato Neurobiologia alla Normale di Pisa ed è direttore dell'Istituto di Neuroscienze del CNR oltre ad avere all'attivo numerose pubblicazioni, il quale mi risulta sia giunto a conclusioni sul valore della funzione della parola e sul campo del linguaggio analoghe a quelle di Ferdinand de Saussure. Cent'anni or sono, nel suo celebre “Corso di linguistica generale” – peraltro molto ben presentato nell'edizione italiana a cura di **Tullio De Mauro** – pone al centro della struttura umana la funzione della parola e della lingua. Distingue il significato come concetto e il significante come immagine acustica: dunque la mente è una macchina di significanti, oltre che di significati. Ricevere significanti dalla radio, dalla televisione, da YouTube o da Spotify non impedisce strutturalmente che si giunga a organizzarli in un significato.

La Biscaldi stessa cita le quattro principali rivoluzioni della comunicazione iniziando con il passaggio dall'oralità alla scrittura. **È nota la posizione di Platone**, espressa 2.500 anni or sono: la scrittura va intesa come farmaco. Farmaco con la duplice accezione, ampiamente sottolineata da Jacques Derrida, di rimedio e di veleno: la scrittura è un rimedio per i limiti della nostra memoria; la scrittura è tuttavia anche un veleno perché ci porta a smarrire il valore della parola orale fino a determinare una “de-menza” – per citare il titolo del testo che riporti e che mi appresto a leggere. L'autrice sembra lucida quando riconosce come i nuovi strumenti di comunicazione convivano con i precedenti, senza sostituirli del tutto: la scrittura convive con l'oralità, i mass media con i libri. Noi potremmo aggiungere che la Play Station convive con i tornei di calcio o di pallavolo, che il sesso online convive con quello offline, che gli incontri sulle app di dating convivono con quelli avvenuti in luoghi tradizionali.

Vi sono senza dubbio **fattori preoccupanti nella diffusione esponenziale dei dispositivi digitali come il cyberbullismo**, il sexting, il revenge porn, l'impatto che i social hanno sulla formazione dell'identità con la conseguenza di favorire lo sviluppo di disturbi alimentari o della cosiddetta disforia di genere.

Spitzer giunge alle stesse posizioni di **Sherry Turkle** sul trovarsi *Insieme ma soli*, titolo di quello che è probabilmente il più celebre libro della collega statunitense. Lo cita esplicitamente. Peccato che la Turkle abbia molto ammorbidito la propria posizione in parte forse anche moralistica, in questi ultimi anni, senza dubbio non senza nessi con la pandemia che ha implicato un upgrade nella

digitalizzazione del mondo per tutti e – ricordiamolo – ha permesso di salvare molte vite umane in quel periodo.

Mantegazza stesso afferma nella suddetta sede che, una volta acquisite le competenze della lettura-scrittura, l'ipertesto digitale può invece costituire una preziosa opportunità di arricchimento tramite cui acquisire ulteriori e più dettagliate informazioni.

Se tutti noi non ricordiamo i numeri di telefono che abbiamo memorizzato nel nostro cellulare, a differenza di quelle che componevamo sulla tastiera del ricevitore di casa, perché non considerare uno sviluppo mentale meno mnemonico e maggiormente volto a cogliere i concetti in quanto tali? Se la diffusione dei sistemi di geolocalizzazione relativi al GPS porta a sviluppare meno la capacità di orientamento quando ci spostiamo per gite e vacanze, perché non capire che vi sarà più tempo da dedicare all'acquisizione di informazioni culturali, storiche, politiche sui luoghi che frequentiamo? Del resto, la dimenticanza non è soltanto un effetto di un deficit neurologico come quello specifico di demenze tipo morbo di Alzheimer ma anche una formazione dell'inconscio. Freud la annovera addirittura come esempio eclatante della psicopatologia della vita quotidiana dedicando a queste forme di oblio il primo capitolo dell'omonimo libro.

Soggetto e struttura clinica

Manfred Spitzer riporta dei fatti di cronaca a riprova delle proprie tesi palesemente critiche nei confronti delle nuove tecnologie. Scrive dei suicidi compiuti per emulazione, dopo aver trascorso del tempo dinanzi al video della TV a vedere un film nel quale si compie un gesto estremo. Auspica, giustamente, un intervento dei mass media nella prevenzione degli agiti autodistruttivi. Lui stesso, però, scrive del celebre "effetto Werther". Come abbiamo studiato al Liceo, dopo la pubblicazione avvenuta nel 1774 del noto e splendido romanzo di Goethe, vi fu un'epidemia di suicidi effetto della suggestione indotta dalla lettura de "**I dolori del giovane Werther**" che narra dell'amore impossibile del protagonista per Charlotte, detta Lotte. L'effetto suggestivo si imprimeva e si imprime tuttora anche attraverso la lettura e portava soggetti strutturalmente volti al passaggio all'atto a suicidarsi.

La violenza fine a sé stessa esiste a prescindere dal digitale

Del resto, sono risapute **le polemiche sui suicidi indotti da brani musicali di alcune band**. In qualità di perito del tribunale di Cottbus, nel Brandeburgo, si è occupato di un omicidio senza apparente motivo. Un giovane, dopo aver giocato tutto il giorno con un amico a un videogame violento, aveva incrociato un cinquantenne senz'atletismo; questi era caduto e il ragazzo lo aveva colpito ripetutamente fino a ucciderlo. Peccato che gli omicidi senza movente siano ampiamente documentati nei trattati della psichiatria francese classica, ben prima che esistessero la televisione e Internet. **Bruce Springsteen** ne canta nel suo celebre brano Nebraska, che fornisce il titolo al suo album del 1982. Racconta l'inquietante vicenda realmente accaduta del diciannovenne Charles Starkweather e della sua fidanzata quattordicenne, Caril Fugate i quali compirono una serie di **assassini senza un movente ben preciso**. In carcere, Charles scrisse poi al papà che non era dispiaciuto per quanto aveva fatto perché perlomeno lui e Caril si erano un po' divertiti. Omicidi immotivati capitano di frequente come tentativo di guarigione in casi di psicosi poiché il soggetto trova o pensa di trovare una pacificazione in certi drammatici passaggi all'atto oppure in quanto si sente paradossalmente finalmente salito al rango di soggetto.

Manca completamente nei testi di Spitzer, così come in “Insieme ma soli” di Sherry Turkle qualsiasi riferimento alla diagnosi differenziale: l'effetto di suggestione indotto dalle immagini degli schermi così come dalla lettura o dall'ascolto di un certo tipo di musica suscita fantasie, pensieri intrusivi, angoscia in soggetti non psicotici mentre **può determinare un passaggio all'atto quasi esclusivamente in soggetti psicotici.**

Affermo inoltre convintamente che ci insegnano maggiormente le esperienze cliniche di quanto ci insegnino ricerche nel campo delle neuroscienze. **Apprendiamo dalla ricerca single case.** A questo proposito, con le giuste accortezze relative alla privacy, possiamo succintamente riferirci a **genitori succubi loro per primi dei dispositivi digitali**, loro stessi da sempre con lo smartphone in mano, dipendenti dalle immagini del tablet. Si tratta forse di soggetti con una forma di debilità (ecco di nuovo la de già incontrata con il termine de-menza), che non riescono a installarsi saldamente nel campo della parola e nel discorso tanto da sopperire grazie agli oggetti digitali alle proprie condizioni di difficoltà relazionale? Se così fosse, l'apprendimento dei loro figli potrebbe venirne evidentemente condizionato a prescindere dall'uso della tecnologia. Si tratta invece di soggetti psicotici, caratterizzati da una forma di psicosi ordinaria, socialmente quasi invisibile, compatibile con l'ordine costituito? Se così fosse, essere figli di genitori psicotici e con carente funzione riflessiva, avrebbe delle ripercussioni sullo sviluppo da valutare accuratamente. Con o senza dispositivi digitali.

Il caso degli hikikomori

Quanto al caso degli adolescenti gamer e dei **cosiddetti hikikomori**, qual è la loro diagnosi differenziale al di là del dato meramente fenomenico? Si tratta forse di una prepsicosi che riesce a intrattenere relazioni soltanto dietro il rassicurante velo dello schermo che fissa la sua libido sull'immagine narcisistica? Sono tutti interrogativi aperti ai quali non possiamo dare risposta in questa sede ma tali da ricordarci il valore della singolarità di ogni posizione soggettiva che va colta nell'intimità dello spazio clinico, della stanza d'analisi. Dietro l'imperante propagarsi degli oggetti digitali in età evolutiva vi sono storie, persone, ragazzi e ragazze con la propria unicità. Le incontriamo come clinici, come docenti, come genitori, come zii, come educatori, come allenatori nelle società sportive.

D'altro canto, i tanto vituperati strumenti digitali risultano spesso preziosi nella pratica clinica, negli ambiti della cura. Si considerino le nuove prassi relative ai trattamenti svolti tramite la cosiddetta realtà virtuale. Riportiamo nello specifico l'esempio di due settori clinici nei quali i dispositivi digitali si stanno dimostrando sempre più importanti e dei quali abbiamo ampiamente scritto nel quarto e ultimo capitolo del volume di 240 pagine “Bucare lo schermo. Psicoanalisi e oggetti digitali”: attacchi di panico e autismo. **Da anni mi capita di ricevere richieste di appuntamenti online da parte di giovani con forme di panico e di agorafobia marcate** i quali non potrebbero mai venire all'appuntamento in presenza; non di rado, dopo un percorso preliminare online, riescono a intraprendere una vera e propria esperienza psicoanalitica in studio. Quanto all'autismo, nel novembre 2019, si è svolto un importante convegno internazionale all'Università di Rennes 2 sull'utilizzo del digitale (*numérique* in francese) e della robotica nella relazione educativa e nel trattamento clinico con ragazzi autistici.

A fianco agli interventi di **Jean-Claude Maleval** del quale ho l'onore di aver tradotto l'unico libro pubblicato in italiano e di Eric Laurent (senza dubbio uno fra i maggiori esponenti della psicoanalisi lacaniana a livello mondiale), vi sono state testimonianze di adolescenti autistici come Théo Fache e di scienziati autistici come il già citato Daniel Tammet. Sottolineavano quanto la relazione regolata dagli oggetti digitali, che possono venire padroneggiati dal bambino autistico, divenga fondamentale per ridurre l'angoscia o la persecutorietà nell'incontro con gli altri sino a permettere

lo sviluppo di competenze, di forme di apprendimento secondo le propensioni e le affinità di ciascuno. Un esempio famoso di *Affinity Therapy* è quello di Owen Suskind, riportato da suo padre Ron il quale vinse a suo tempo un premio Pulitzer per la propria attività giornalistica: dai 3 ai 9 anni, non parlava; trascorreva gran parte delle proprie giornate dinanzi allo schermo per vedere cartoni animati di **Walt Disney finché il papà capì che i “versi” emessi da Owen erano versioni storpiate delle parole dei cartoon**. Da lì uno sviluppo soggettivo importante fino al punto di arrivare a 20 anni a poter parlare in conferenze e in contesti universitari come appunto quello di Rennes, la maggiore realtà universitaria della Francia dopo Parigi. “*Life animated*” è il film che narra la vicenda Suskind.

Conclusioni

In ultima istanza, approcciare gli interrogativi sollevati dalla diffusione imperante dei dispositivi e degli oggetti digitali, nell’epoca nella quale gli oggetti sopravanzano per rilevanza gli ideali tradizionali, ci impone di considerare la peculiarità di ogni posizione soggettiva. Per motivi etici e per problemi clinici, la peggiore fra le operazioni che si potrebbero attuare sarebbe quella di **fare di tutta tutta l’erba un fascio**. La nostra tesi fondamentale è che vada sottolineata la funzione, di crescita o di stagnazione, che l’oggetto digitale determina su un soggetto in base alle sue proprie caratteristiche e alla specificità del momento attraversato. Anche e soprattutto in età evolutiva.

Bibliografia

Biscaldi A. – Matera V. (2019), *Antropologia dei social media. Comunicare nel mondo globale*, Carocci, Roma.

Christakis D. - Zimmerman F. - Di Giuseppe D. L. - Mc Carthy C. (2004), *Early television exposure and subsequent attentional problems in children*, *Pediatrics* 113, pp. 708 . 713.

Derrida J. (1968), *La farmacia di Platone*, Jaca Book, Milano, 2015.

Ennemoser M. – Schneider V. (2007), *Relations of television viewing and reading: findings from a 4-year longitudinal study*, *Journal of Educational Psychology*, 99, pp. 349-368.

Freud S. (1901), *Psicopatologia della vita quotidiana*, Opere, Bollati Boringhieri, Torino, 1989.

Goethe W. (1774), *I dolori del giovane Werther*, Universale Economica Feltrinelli, Milano, 2014.

Maffei L. (2018), *Elogio della parola*, Il Mulino, Bologna.

Maleval J.-C. (1991), *Logique du meurtre immotivé* in « *Psychose naissante, psychose unique ?* » opera collettiva sotto la direzione di Grivois H., Masson, Parigi, Milano, pp. 43 – 67.

Maleval J. – C. (2011), *Isteria e follia. Logica del delirio come tentativo di guarigione*, Bruno Mondadori, Milano.

Perrin M. (2015), *Affinity therapy. Nouvelles recherches sur l’autisme*, Presses Universitaires de Rennes.

Pozzetti R. (2021), *Bucare lo schermo. Psicoanalisi e oggetti digitali*, Alpes, Roma.

De Saussure F. (1916), *Corso di linguistica generale*, Laterza, Bari, 1967.

Sisask M. – Varnik A. (2012), *Media roles in suicide prevention: a systematic review*, International Journal of Environmental Research and Public Health, 9, 123: 138.

Suskind R. (2014), *Reaching my autistic son through Disney*, The New York Times, 9/3/2014.

Turkle S. (2011), *Insieme ma soli. Perché ci aspettiamo sempre più dalla tecnologia e sempre meno dagli altri*, Codice, Torino, 2012.

A. Cangini, *Sull'impatto del digitale sugli studenti, con particolare riferimento ai processi di apprendimento*, Documento della Commissione Parlamentare (Istruzione Pubblica, beni culturali), seduta del 9 giugno 2021. [↑](#)



L'interpretazione dei dati: il processo di riappropriazione in quattro fasi

Dare significato a dati che di per sé non ne posseggono vuol dire trasformare i dati in significati. E cioè dare valore a ciò che, nelle mani dei singoli, non sarebbe altro che una potenzialità inutile. Ecco come riappropriarsi dei dati

Di **Maurizio Ferraris**, professore ordinario di filosofia teoretica presso la Facoltà di Lettere e Filosofia dell'Università degli Studi di Torino

L'**interpretazione dei dati** si impone come esigenza di un'**ermeneutica digitale**, compiuta da corpi intermedi, capaci di dare valore a ciò che, altrimenti, nelle mani dei singoli, non sarebbe altro che una potenzialità inutile. Anche perché la produzione di valore più significativa è quella ottenuta dall'aggregazione e dalla correlazione di masse di dati quanto più ampie possibili.

Se una struttura di intermediazione dispone dei dati, è necessaria un'interpretazione che dia significato a dati che di per sé non ne posseggono.

Trasformare i dati in significati costituisce dunque l'enorme sfida intellettuale che finora si è svolta nel segreto delle piattaforme cinesi e americane. Invece ora è tempo di riappropriarsi dei dati. Ecco come, in quattro passaggi.

La riforma protestante e il neoliberismo

La riforma protestante ha provocato il **risveglio dell'ermeneutica biblica**. Il cuore della presa di posizione protestante consiste infatti nel riconoscere l'essenza della fede non nell'accettazione della tradizione apostolica romana, ma nel contatto diretto del credente con il testo biblico.

Questo però comporta la costituzione di canoni ermeneutici che possano guidare il credente nell'accesso al testo sacro.

La spinta emancipativa della Riforma era minacciata da due rischi potenziali, che effettivamente si manifestarono nel corso del tempo. Il primo è la **personalizzazione della fede e la crescita di un atteggiamento individualistico** che costituisce l'antefatto generale del cosiddetto "**neoliberismo**".

Il secondo è che **nessun prontuario ermeneutico può avanzare la pretesa di costituire, per l'utente, una guida certa e infallibile**. Cioè, tradotto dei termini del problema che qui interessa, **nulla garantisce che un canone di ermeneutica dei dati possa da sé solo offrire una alternativa alla grande esperienza**, e alla potenza di calcolo, delle piattaforme commerciali capitaliste o delle piattaforme statali comuniste.

Ancor più, per quanto riguarda l'individualismo, occorre non dimenticare che **la natura stessa della materia sottoposta a interpretazione dei dati è tale che il valore dei dati di un singolo è**

raramente rilevante, laddove la produzione di valore più significativa è quella che ha luogo attraverso l'aggregazione e la correlazione di masse di dati quanto più estese possibili.

L'interpretazione dei dati e la rivoluzione marxista

In questo quadro, il problema ermeneutico della riforma si intreccia con il problema sociale della rivoluzione marxista. L'interpretazione della Bibbia è comprensione di un testo la cui ispirazione si presenta come soprannaturale.

L'interpretazione dei dati è invece il tentativo di dare significato al prodotto di una attività umana. Da questo punto di vista, la situazione con cui si trova a misurarsi l'ermeneutica digitale appare molto vicina a quella con cui si è misurato Marx nella analisi del plusvalore.

Se il servo della gleba sapeva esattamente quando cessava di lavorare per sé e iniziava a lavorare per il padrone, giacché gli erano affidati due campi, il suo e quello del feudatario, l'operaio in fabbrica non saprà mai in quale momento preciso della sua giornata di lavoro cessa di pagarsi lo stipendio e inizia a lavorare a beneficio esclusivo del capitalista.

La situazione è tanto più complessa in quanto neppure il capitalista sarebbe in grado di rispondere a questo interrogativo, giacché a deciderlo sono le condizioni del mercato, sicché l'unica cosa certa è che, qualora il numero delle ore lavorate bastasse esclusivamente a coprire il salario del lavoratore, la fabbrica chiuderebbe.

Produzione di valore: la differenza fra Cina e Stati liberali

Nella condizione attuale, si viene a creare la situazione per cui potenzialmente l'intera umanità, e attualmente una larga parte di essa, anziani, bambini, disoccupati, altrimenti occupati, produce valore potenziale a vantaggio esclusivo delle piattaforme. Beninteso, queste possono anche decidere di socializzare gli utili, come avviene in Cina, ma ciò al prezzo di una totale perdita della libertà individuale.

La soluzione di compromesso a cui sono giunti gli Stati liberali è una forma **moderata di welfare ottenuta attraverso la tassazione delle piattaforme**.

Tuttavia, questa tassazione è difficile da esercitarsi perché avviene in un contesto di opacità cognitiva, caratterizzata dall'assenza di un mercato che determini pubblicamente il valore dei dati, e dalla conseguente impossibilità di determinare quali siano gli utili effettivi delle piattaforme. Più in generale, il ricorso alla tassazione, che era lo strumento principe del welfare tradizionale, appare altamente problematico non solo perché coloro che attualmente realizzano i maggiori utili sono anche i più difficili da tassare equamente, mentre i soggetti più facilmente tassabili sono economicamente molto più deboli.

Ma soprattutto per un motivo concettuale. E cioè che il welfare tradizionale si limita alla redistribuzione più equa di valori già esistenti, ossia non intercetta la novità è il carattere proprio del patrimonio dell'umanità, il fatto cioè che sia un capitale totalmente nuovo e senza precedenti nella storia del mondo, reso possibile dall'apparizione del digitale come apparato universale di registrazione.

Webfare e welfare a confronto

Per cogliere la specificità dell'apparizione del digitale come apparato universale di registrazione e le caratteristiche distintive del webfare rispetto al welfare, conviene anzitutto far ritorno all'analogia con il plusvalore che abbiamo fatto poco sopra.

Rispetto al rapporto tra lavoro salariato e capitale, la mobilitazione degli umani sul web presenta due significative differenze. La prima è che, diversamente che nelle fabbriche tradizionali, nel caso delle piattaforme il lavoro è erogato in forma volontaria, e non comporta né costrizione né alienazione, il che genera l'apparenza di uno scambio equo tra utente e piattaforma, in cui l'utente fornisce gratuitamente informazioni e servizi, in cambio di informazioni e servizi gratuiti erogati dalla piattaforma.

Si tratta soltanto di un'apparenza, giacché la piattaforma può ricavare un valore molto maggiore dai dati di quanto non ne possa ottenere l'utente, appunto perché possiede gli **strumenti concettuali e tecnologici necessari per interpretare un capitale di dati a cui l'utente non ha accesso, perché non sono visibili nell'infosfera, ma sono prodotti nella docusfera, ossia sono metadati, e soprattutto sono dati che le piattaforme aggregano**, mentre nella migliore ipotesi, **avvalendosi della legge per la portabilità dei dati, l'utente può al massimo avere accesso ai propri dati individuali, con valore prossimo a zero.**

Sono proprio queste circostanze che rendono cruciale il problema di una **ermeneutica dei dati** compiuta da **corpi intermedi, capaci di dare valore a ciò che, altrimenti, nelle mani dei singoli, non sarebbe altro che un'inutile potenzialità.**

La metafora dei dati come nuovo petrolio è infondata

Da questo punto di vista, **la metafora dei dati come nuovo petrolio è ingannevole per più di un motivo.**

Innanzitutto, **i dati, diversamente dal petrolio, sono rinnovabili.** In secondo luogo, i dati, diversamente dal petrolio, **sono il risultato di attività umane registrate sul web.** Infine, diversamente dal petrolio, **i dati non richiedono apparati onerosi per l'estrazione, la raffinazione e la distribuzione.**

La prima e la seconda circostanza rendono in linea di principio attuabile l'uso dei big data per finalità mutualistiche e solidali, dal momento che la legislazione europea consente ai singoli di richiedere alle piattaforme i loro dati, ciò che **da parte dei singoli si configura come una richiesta lecita giacché sono loro i produttori di loro dati**, e dal punto di vista delle piattaforme non appare come una sottrazione illecita giacché **le piattaforme non si privano dei dati che condividono con gli utenti.**

A questo punto, se una struttura di intermediazione dispone dei dati, **il solo ostacolo è costituito dalla interpretazione che dia significato a dati che di per sé non ne posseggono.** Trasformare i dati in significati costituisce dunque **l'enorme sfida intellettuale che per il momento si è svolta nel segreto delle piattaforme cinesi e americane.**

Questo insieme di canoni e di esempi ermeneutici è parte integrante del patrimonio dell'umanità, giacché permette di valorizzarlo, e come tale deve essere pubblico.

L'ermeneutica digitale e l'interpretazione dei dati

In questo quadro, l'ermeneutica digitale è produzione di significati attraverso i dati. E i significati sono portatori di valore per utenti umani. Tra i valori, il più universale e pregiato è quello della verità, per gli evidenti vantaggi pratici che può portare all'umanità.

Questo però significa che, diversamente dalla realtà, che è una dimensione ontologica che riguarda tutto ciò che esiste indipendentemente dalla conoscenza che ne abbiamo, **la verità è una dimensione epistemologica sorretta da un principio antropico debole.**

Con questa espressione intendo dire che **la verità è interessante e significativa solo nel quadro della forma di vita umana (principio antropico)**, ma che questo non fa della verità qualcosa di soggettivo, né di semplicemente dipendente da un accordo intersoggettivo (ed è per questo che il principio antropico è debole).

Le metodiche dell'interpretazione dei dati

Occorre determinare la natura delle istituzioni preposte alla capitalizzazione alternativa dei dati; le modalità specifiche attraverso cui si attua la capitalizzazione; le forme di socializzazione dei valori ottenuti, sia in forma monetaria sia in termini di servizi e garanzie.

Tutto ciò però richiede una condizione preliminare, e cioè che si determinino delle metodiche di interpretazione dei dati che siano alla portata tecnica di istituzioni che non dispongono di meccanismi di calcolo ultra potenti, e che riescano a intercettare in maniera originale rispetto alle metodiche delle grandi piattaforme dei criteri di valorizzazione, rilevanza e utilità dei dati.

Considerando che **le istituzioni mediatrici che si faranno carico di richiedere alle piattaforme commerciali i dati degli utenti**, di cui avranno ottenuto il consenso, **saranno in larga misura istituzioni pubbliche, università, ospedali, banche**, va tuttavia tenuto presente, come criterio fondamentale, che queste istituzioni, diversamente dalle piattaforme commerciali, dispongono già di ampie banche dati ordinate, il che di per sé facilita e orienta l'opera di interpretazione, permettendo in particolare di aver chiari gli scopi e i bisogni a cui l'interpretazione dei dati deve saper rispondere.

La riappropriazione dei dati in 4 fasi

La logica di questa riappropriazione - che non è esproprio - poggia sulla circostanza per cui i dati, proprio come le idee, si possono condividere e riusare quanto si vuole. E richiedere i dati alle piattaforme commerciali non significa chiedere che non li usino più, tutt'altro. Ciò previene l'obiezione: come sia possibile convincere le piattaforme commerciali a condividere la loro ricchezza? Semplice. Non c'è alcun bisogno di convincerle, visto che **da loro non si tratta di pretendere soldi, ma dati, quelli che noi stessi produciamo.**

Le **piattaforme rimangono in possesso dei dati**, oltre che dei proventi delle loro capitalizzazioni, **il che non esclude la possibilità di capitalizzazioni alternative dei dati da parte di soggetti diversi dalle piattaforme.**

Se, con un termine enfatico ma, mi auguro, espressivo, ho battezzato queste agenzie di capitalizzazione alternativa **"banche della virtù"**, è perché si caratterizzano per la capitalizzazione dei dati prodotti da tutti coloro che liberamente avranno aderito al loro progetto ("da ognuno secondo le sue capacità") e per la redistribuzione dei proventi della capitalizzazione a coloro che, tra

coloro che aderiscono al patto di capitalizzazione, non possiedono soldi o altre risorse, ma semplicemente dati (“a ognuno secondo i suoi bisogni”).

Questa circostanza previene un'altra possibile obiezione: c'è un problema etico dietro questa alleanza con le piattaforme che basano i loro modelli di business sull'elusione fiscale, i lavoretti di persone sottopagate e una dipendenza dai mezzi tecnologici? Ora **non c'è alcuna alleanza con le piattaforme commerciali**. Semplicemente, in base alla legislazione europea, gli utenti si fanno dare i dati che producono, e decidono di investirli usando delle altre piattaforme virtuose. **La riappropriazione avviene attraverso quattro operazioni: qualificazione, quantificazione, compensazione e condivisione.**

La qualificazione

Il primo passo, l'acquisizione dei dati, è la loro qualificazione, ossia **il riconoscimento del fatto che il loro valore dipende in pari misura dalla mobilitazione dell'umanità** e dal fatto che questa mobilitazione è intercettata e interpretata dalle piattaforme. Quanto abbiamo detto sin qui costituisce un argomento sufficiente per fondare il diritto di acquisizione dei dati, che personalmente preferisco giustificare in base alla prestazione d'opera (la mobilitazione) che gli umani esercitano sulle piattaforme piuttosto che con il richiamo all'esercizio di un diritto di proprietà.

I dati in quanto tali, infatti, non esisterebbero senza le piattaforme e si trovano nelle piattaforme, che ne sono proprietarie; ciò rende dubbia la legittimità della richiesta di una proprietà libera dei dati a cui farebbe seguito una vendita. Visto che però le piattaforme non sono proprietarie della mobilitazione che produce i dati, è proprio attraverso il richiamo a quest'ultima che si può fondare la richiesta di condivisione in vista di una capitalizzazione alternativa.

La quantificazione

Il secondo passo è la quantificazione. Occorre trovare dei metodi per capire quanti dati produciamo nel nostro rapporto con le piattaforme. Esistono già proposte e iniziative (Lehdonvitra et al 2016 ^[5]) volte a chiedere il controllo sui propri dati (Tirole 2017 ^[9]) e a quantificarli (Bolognini e de Michelis 2018 ^[2]).

A tal fine, abbiamo a che fare con un caso di serendipity: le norme elaborate in sede europea per la tutela della privacy, che poggiavano su un fondo di civil law, l'inalienabilità dei diritti della persona, divengono lo strumento per riconoscere la quantità di dati prodotti.

Nell'ambito della quantificazione dei dati è così possibile formulare una legge: quanta minore è la privacy, tanto maggiore è il numero di dati raccolti dalle piattaforme, e inversamente. Ma quantificare i dati non è ancora quantificare il valore dei dati, e questo è il grande problema. Contrariamente alle letture che vedono nell'economia digitale un prevalere del mercato sull'azienda (Mayer-Schönberger e T. Ramge 2019 ^[7]), qui l'azienda si sostituisce al mercato e rende impossibile la determinazione di un valore pubblico.

Quanto valgono le informazioni che permettono di far volare gli aerei a pieno carico grazie ai dati sui comportamenti dei passeggeri? Ecco una domanda a cui non si potrà mai rispondere se l'acquisto dei dati è il frutto di una trattativa privata tra una piattaforma e una compagnia. Se esistesse un mercato (e non potrà non esistere con l'ingresso di investitori di dati che siano diversi

dalle piattaforme) ci sarebbero una domanda e un'offerta e, su quella base, si determinerebbe il valore.

La compensazione

È questa difficoltà, cioè in ultima analisi l'assenza di una borsa dei dati, che rende difficile l'attuazione di politiche di compensazione volte a redistribuire il plusvalore delle piattaforme commerciali.

Questo vale, prima di tutto, per le proposte di tassazione (Bolognini e de Michelis 2018; Montes, Sand-Zantman e Valletti 2019 ^[8]) prospettate dalla Cina (Matsuda 2021 ^[6]) e già in parte attuate dagli Stati Uniti e dalla Unione Europea ^[1].

Contrariamente a quanto suggerito da alcuni, queste iniziative non corrono il rischio di far ricadere i costi sugli utenti (Casilli e Bouquin 2020; Casilli 2020; Lehdonvitra et al. 2016 ^[4]).

Le piattaforme perderebbero tutta la loro attrattività qualora cessassero di fornire gratuitamente dei servizi, ma trovano il loro limite maggiore proprio nella circostanza per cui, sino a che non ci sarà una borsa dei dati, risulta molto difficile determinarne il valore, e dunque esercitare la giusta pressione fiscale sulle piattaforme.

La compensazione, tuttavia, non si limita alla redistribuzione del prelievo fiscale, e in particolare (valorizzando il concetto di "patrimonio dell'umanità") può consistere nell'accrescimento delle zone di gratuità di beni e servizi, del resto coerente con gli interessi commerciali delle piattaforme. Coerentemente con l'impostazione generale che sto seguendo, secondo cui l'apporto degli utenti consiste nella mobilitazione, e quello delle piattaforme nella registrazione e dunque nella produzione dei dati, la gratuità non deve essere considerata come un bene comune, ma piuttosto come un prodotto cooperativo. Con questa differenziazione terminologica indico la circostanza per cui attraverso la gratuità gli utenti non vedono riconosciuto un diritto sul patrimonio ottenuto dalla capitalizzazione dei loro dati da parte delle piattaforme (esercitare prelievi sul plusvalore e ridistribuirli è di competenza della fiscalità statale), bensì l'apporto della loro mobilitazione nella produzione dei dati.

La condivisione

Il quarto passo, diverso dalla compensazione e che richiede l'intervento della banca della virtù è la condivisione non dei dati, ma del loro valore. Se la compensazione consiste in un intervento su dati già capitalizzati, qui abbiamo a che fare con due processi radicalmente differenti. In primo luogo, abbiamo a che fare con una produzione di valore attraverso una capitalizzazione alternativa e autonoma rispetto a quella attuata dalle piattaforme commerciali, il cui apporto si limita alla condivisione dei dati con le banche della virtù.

In secondo luogo, affinché il termine "banca della virtù" non appaia abusivo e insopportabilmente retorico, è necessario che la virtù sia effettiva, e cioè che sia chiaro che lo scopo della banca della virtù non consiste nel remunerare i correntisti (per questo si deve ricorrere ai servizi tradizionali della banca e della borsa, che valgono per chi ha soldi), bensì far entrare nel gioco economico quella stragrande maggioranza dell'umanità che non ha soldi ma ha dati, e che deve essere inserita nel mondo della cittadinanza non formale ma sostanziale attraverso l'apertura di un conto in banca fatto prima di dati, poi di soldi. Ma perché questo obiettivo possa venire conseguito occorre

anzitutto che la banca della virtù si doti di strumenti politici che qualifichino i suoi correntisti come produttori di valore.

Bibliografia

Per esempio, con la Digital Tax. [↑](#)

Bolognini L. e I. de Michelis (2018), “An Introduction to The Right to Monetize”, (RTM), 9 April 2018. [↑](#)

Casilli, A. (2020). Schiavi del clic (2019), tr. it. Feltrinelli, Milano. [↑](#)

Casilli, A. e S. Bouquin (2020). “Il n’y a pas d’automatisation sans micro-travail humain”, Les Mondes du Travail, pp. 24-25. [↑](#)

Lehdonvitra, A. et al. (2016). Data Financing for Global Good: A Feasibility Study, Oxford Internet Institute, Oxford. [↑](#)

Matsuda, N. (2021) <https://asia.nikkei.com/Economy/Is-China-considering-a-data-tax-on-big-tech-Signs-point-to-yes>, November 22, 2021. [↑](#)

Mayer-Schönberger, V. e T. Ramge (2019). Reinventing Capitalism in the Age of Big Data, Hodder and Stoughton, London. [↑](#)

Montes, R. W. Sand-Zantman e T. Valletti (2019). “The Value of Personal Information in Online Markets with Endogenous Privacy”, Management Science, INFORMS, 65, 3, March 2019, pp. 1342-1362.. [↑](#)

Tirole, J. (2017). Economics for the common good, Princeton University Press, Cambridge (MA). [↑](#)

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

©ICT & Strategy