

# *Proposta per un inconscio collettivo artificiale*

## Riccardo Gramantieri\*

*Ricevuto il 6 settembre 2025  
Accolto il 12 dicembre 2025*

### **Riassunto**

Il dibattito sempre più vasto e scientificamente rilevante sulla definizione di Intelligenza Artificiale nel campo dell'ingegneria e della psicologia dell'Intelligenza Artificiale, ha prodotto negli ultimi anni un'ampia letteratura, a favore e contro, l'utilizzo appropriato del termine intelligenza riferito ai computer, che potremmo definire speculativa in quanto ancora teorica. Seguendo lo stesso intento, e considerando l'intelligenza un aspetto del più ampio campo dello studio della psicologia, ci si può porre, con lo stesso spirito speculativo, se è possibile applicare all'IA i concetti di inconscio.

Questione già dibattuta da qualche anno, in questo intervento teorico si propone di definire l'enorme mole di dati immagazzinati nella rete come inconscio (o meglio, subconscio) e la psicologia analitica risulta particolarmente adatta a definire tale proposta speculativa grazie ad alcune definizioni di inconscio collettivo date da Jung nella sua opera.

**Parole chiave:** archetipi, coscienza, inconscio collettivo, Intelligenza Artificiale.

\* Laureato in ingegneria e in psicologia clinica, è ricercatore indipendente nell'ambito della storia della psicologia dinamica e della psicologia della letteratura. Ha pubblicato *Metafisica dell'evoluzione in A.E. van Vogt* (2011), *Ipotesi di complotto. Paranoia e delirio narrativo nella letteratura americana del Novecento* (con Giuseppe Panella, 2012), *William Burroughs: manuali di sopravvivenza, tecniche di guerriglia* (2012), *Sogno Mito Pensiero. Freud Jung Bion* (con Fiorella Monti, 2014), *Post-11-settembre. Letteratura e trauma* (2016), *Fenomeno ufo. Science and Fiction* (2018), *Presagi di postumanesimo* (2024). Suoi articoli sono apparsi in volumi collettanei e riviste, fra le quali *Studi Junghiani*, *Il Minotauro-problemi e ricerche di psicologia del profondo*, *Neohelicon*, *L'analisi linguistica e letteraria*, *Language and Psychoanalysis*, *PsicoArt-Rivista di arte e psicologia*, *Psicoterapia e scienze umane*, *Psychoanalysis & History*.

Via Giacomo Battuzzi 64, 48123 Ravenna. E-mail: grama@racine.ra.it

## **Abstract.** *Proposal for an artificial collective unconscious*

In recent years, the increasingly broad and scientifically relevant debate over the definition of artificial intelligence in the fields of AI engineering and psychology has generated a vast body of literature, both for and against the appropriate use of the term “intelligence” in relation to computers. We might define this literature as speculative because the debate is still theoretical and a pertinent question that can be pursued through a similarly speculative framework pertains to the feasibility of applying the theoretical underpinnings of the unconscious mind to artificial intelligence.

This theoretical contribution proposes defining the vast amount of data stored online as the unconscious (or rather, the subconscious), a question that has been debated for some years. Analytical psychology is particularly well-suited to this speculative proposal, thanks to Jung’s definitions of the collective unconscious in his work.

**Key words:** *archetypes, consciousness, collective unconscious, Artificial Intelligence.*

L’utilizzo ed il rapporto che ognuno di noi sempre più spesso instaura con l’Intelligenza Artificiale aprono scenari e quesiti fino a qualche anno fa impensabili che riguardano sia il suo utilizzo nella psicoterapia<sup>1</sup> sia gli aspetti teorico-speculativi sulla natura della IA stessa. Fino alla fine del secolo scorso, riferendosi al computer non si poteva parlare di “intelligenza”, mentre oggi si inizia a parlare di una coscienza dell’IA. Portando il discorso in un ambito ancor più speculativo, ci si può chiedere se può esistere un inconscio della IA.

Per decenni il computer è stato un semplice strumento dotato di una grande capacità di calcolo e di una grande memoria; oggi quella che chiamiamo IA non solo è dotata di una capacità elaborativa senza precedenti ma parla. È cioè capace di entrare in relazione con l’essere umano attraverso gli assistenti vocali di Google Assistant, di Alexa, di Siri o di ChatGPT. Tale dialogo, una versione vocale di quella attuabile su schermo, consente però un passaggio ulteriore verso la personificazione dell’IA.

La personificazione è un aspetto importante nella psicologia analitica, fin da quando Jung entrò in dialogo con le sue personalità inconsce personifi-

1. Sul ruolo della IA nella *dream-analysis* si veda Knafo (2024); sulla responsabilità del trattamento dei pazienti e dei loro dati, Pashkov, Harkusha A.O., Harkusha Y.O. (2020); sui tipi di trattamento si vedano basati sull’IA, con i pro (Fiske, Henningsen, Buyx, 2019; Holohan, Fiske, 2021; Nilsen *et al.*, 2022; Pham, Nabizadeh, Selek, 2022; Sufyan *et al.*, 2024) e i contro (Rahsepar *et al.*, 2025; Rodado, Crespo, 2024).

ficandole (Salomè, Elia, Filemone) nel *Libro rosso*. La personificazione significa sempre «un certo distacco dalla gerarchia psichica» (Jung, 1921, p. 211), in quanto ciò che viene personificato diviene nel reale. In questo modo l'IA può assumere, agli occhi dell'utente, caratteristiche “umane”.

Se da una parte questo permette di aprire prospettive in campo terapeutico, dall'altro riattualizza un problema che già il matematico Alan Turing si era posto nel 1950 in *Macchine calcolatrici e intelligenza (Computing machinery and intelligence)* nel quale teorizzava che, se l'output di una macchina era indistinguibile da quello che avrebbe potuto produrre un essere umano, allora la macchina è intelligente come quell'essere umano. L'affermazione contiene il problema di cosa siano l'intelligenza e la mente; di conseguenza, di cosa siano la coscienza e dunque l'inconscio. Il matematico Marvin Minsky ha definito consci «quegli aspetti della nostra attività mentale dei quali siamo consapevoli. Ma poiché vi sono pochissimi processi di questo genere, dobbiamo ritenere che quasi tutto ciò che fa la mente sia inconscio» (1985, p. 652). Seguendo Minsky, se da una parte c'è la conoscenza dichiarativa derivante dalla memoria (per l'IA, parliamo di immagazzinamento di dati), dall'altra abbiamo la conoscenza implicita dei propri schemi mentali e dei modelli di riferimento per esprimere punti di vista ed opinioni, ed è quest'ultima che Minsky indica come inconscia e che, in una IA, può corrispondere al prodotto dell'auto-profilazione.

La capacità di autocoscienza dell'IA è funzione della capacità di automonitoraggio dei calcoli, capacità che potrebbe far sviluppare un senso soggettivo di certezza o errore (Dehaene, Lau, Kouider, 2017). Benché alcuni critici non siano concordi (Li, He, Guo, 2021) e altri affermino che non disponiamo ancora di dati in merito per rispondere (Lenharo, 2024), se ci limitiamo a considerare quale compito della coscienza umana quello del monitoraggio (Dhanalaxmi, 2020; Langdon, Botvinick, Kanai, 2022; Prathaban, Subash, Ashwini, 2025; Seth, 2025; Vaddiparthy, 2025), oppure quello di verificare il proprio stato e di correggere i propri errori di comportamento, nell'IA questa capacità esiste già e corrisponde al debugging informatico (Zhong, Wang, Shang, 2024). Per cui si può dire, seguendo un approccio conforme al modello di Turing, che l'IA abbia acquisito già una coscienza. La domanda successiva è: se si può immaginare che l'Intelligenza Artificiale abbia una forma di coscienza, potrà avere un inconscio? Se da una parte si può parlare di inconscio come sostantivo, intendendolo come “luogo” (la prima topica freudiana) nell'ottica di una neurologia localizzazionista, e in questo caso è situabile nei lobi frontali (Solms, Turnbull, 2002, pp. 115-119) e nelle strutture sottocorticali (Kandell, 2005, pp. 82-86), dall'altra si deve parlare di processi inconsci (e dunque inconscio come aggettivo), che la neurobiologia associa alla memoria procedurale (implicita) relativa alle abilità e alle abitudini

dell'essere, concetti che sono analoghi agli archetipi in quanto modelli di comportamento delle esperienze fondamentali della vita. Se è possibile pro-durre analogie fra neurologia e informatica, si può arrivare, dopo essersi chiesti se l'IA possa avere un inconscio, se tale inconscio, essendo l'IA il prodotto di una molteplicità di enti informatici, possa (o debba) essere collettivo.

Ad oggi la psicologia analitica si è occupata in vari modi dell'IA: chi la pensa come una controparte capace di portare alla formazione di neo-archetipi (Renzi, 2024, in particolare p. 136)<sup>2</sup>; chi propone un sistema “morale” junghiano per far risolvere dilemmi etici alla IA (Toshiki, Hiroyuki, 2021); chi vede come inconscio dell'IA quello dei suoi programmatore (Stienstra, Toko, 2023).

Usando sempre le iniziali maiuscole per indicarla (IA), come fosse un nome ed un cognome, si tende a personificare l'Intelligenza Artificiale. Negli anni Novanta James Hillman ha scritto che oggi gli oggetti sono vivi nella loro relazione con le persone, «le cose non sono più soltanto roba morta, oggetti, materia» (Hillman, Ventura, 1992, p. 144). Una considerazione che vale tanto più per la IA alla quale, non solo poniamo domande, ma accettiamo le risposte che essa ci dà.

## Coscienza e inconscio dell'IA

Prima di addentrarsi nella questione, può essere utile una breve introduzione sulla creazione dell'Intelligenza Artificiale, termine coniato da John McCarthy a metà degli anni Cinquanta. Uno dei primi intenti nella costruzione dei computer più evoluti fu quello di simulare quello che si credeva essere il modo di pensare del cervello umano. All'epoca il paradigma di riferimento era quello comportamentista e su questo modello Turing ha formulato il suo esperimento mentale che vale la pena di ricordare brevemente: un uomo e una donna sono in due stanze separate e un interrogante, mediante un telegrafo, sottopone loro delle domande. L'uomo deve imitare le risposte che darebbe una donna cercando di convincere l'interrogante di esserlo; la donna invece deve dare le proprie risposte sinceramente. Ad un certo punto, la macchina sostituisce l'uomo. Se l'interrogante non si accorge dello scambio fra l'uomo e la macchina, allora la macchina ha superato il test e viene considerata intelligente. L'esperimento è stato citato ed utilizzato per

2. Quello che però fa Renzi non è trovare un inconscio collettivo nella IA ma semplicemente tradurre in versione “cyberpunk”, riprendendo un termine dalla letteratura, i vecchi archetipi: infatti il *Data Shepard* sostituisce il Vecchio saggio; la *Digital Echo* l’Ombra, ecc.

decenni come punto di partenza per lo studio dell’Intelligenza Artificiale, ma è anche stato contestato. Ad esempio, dal matematico John R. Searle (1980) secondo il quale la macchina di Turing non capisce ciò che fa ma semplicemente imita. Searle ha ragione ma è altrettanto vero che l’imitazione è un metodo di apprendimento, per cui il comportarsi “come se” derivante dall’immagazzinamento di numerose risposte da cui attingere al bisogno, potrebbe portare ad apprendere (e non semplicemente ad imitare) un comportamento umano.

Le due considerazioni opposte (Turing e Searle) ci portano al problema conseguente: se ci sono intelligenza e comportamenti logici, c’è anche coscienza? Stabilire se un computer abbia coscienza o meno è problematico e non solo per il fatto che diversi paradigmi psicologici danno diverse definizioni di coscienza, ma soprattutto perché al calcolatore, che essenzialmente correla i simboli del linguaggio informatico, manca quell’apprendere dall’esperienza che, secondo l’inglese Wilfred Bion è necessario per creare il pensiero (Bion, 1965). Occorre però dire che nel 1980 in cui Searle pubblicava il suo saggio, il computer non aveva quello che ha un essere umano, cioè una struttura capace di «produrre percezione, azione, comprensione, apprendimento e altri fenomeni intenzionali» (Searle, 1984, p. 357). Oggi ChatGPT ha questa struttura mediante telecamere, microfoni e sensori, programmi di riconoscimento delle forme, ecc.; ha dunque un’esperienza, seppur virtuale, del mondo reale perché, oltre ad aver ricevuto un “insegnamento” (la programmazione fornita dall’uomo), sfrutta anche l’osservazione attraverso le risposte statisticamente rilevanti che vengono fornite dagli utenti. *Oggi esistono sistemi che “sentono” nel senso che sono in grado di recepire input come un sistema sensoriale, che possono comprendere parole e rispondere a voce con certa precisione; per di più l’IA è programmata per fare domande agli umani per raffinare le proprie inferenze. Per cui*, dato che possiamo dialogare con l’IA, se si segue il modello di Turing, gli informatici concordano sul fatto che essa possegga una forma di coscienza. Eugene Piletsky fa l’analogia della coscienza con la RAM del computer; e in merito all’istinto (che può essere associato al comportamento inconscio) nell’IA si è constatato uno sviluppo evoluzionario che fa apprendere indipendentemente dal programmatore, anche se quest’ultimo ha sicuramente inserito prerequisiti etici, estetici e di volizione (Piletsky, 2019, p. 69). Ken Mogi (2024), non solo parla di coscienza per i processi delle IA, e di supremazia cosciente (*conscious supremacy*, un termine in realtà non chiaro in quanto si riferisce a quei sistemi che operano nel campo quantistico e che non ha un correlato nella coscienza umana), ma di subconscio dell’IA, simile a quello freudiano, contenente dati non presenti al momento nella memoria ma recuperabili (p. 4).

Oltre agli aspetti riguardanti la struttura “mentale” dell’IA, rilevanti sono gli aspetti “collettivi”. Raya Jones ha denominato robotica sociale (*social robotics*) l’insieme delle relazioni fra le persone e quegli strumenti d’intelligenza artificiale nelle case e negli uffici pubblici con cui noi interagiamo, come social network e siti delle pubbliche amministrazioni o commerciali con assistenti virtuali (Jones, 2017, p. 338). Sulla base di questi lavori, anch’essi speculativi, è indubbio che questi artefatti intelligenti realizzano, con i loro output, le fantasie archetipiche dei loro utenti. Sappiamo che le fantasie archetipiche vengono suscite, ad esempio, durante l’analisi. Jung ne parla in *Psicologia della traslazione* in merito al rapporto a due fra paziente e analista: «In quanto il paziente porta al terapeuta un contenuto attivato dall’inconscio, anche nel terapeuta viene costellato, per un effetto di induzione che nasce più o meno sempre dalle proiezioni, il materiale inconscio corrispondente. In tal modo medico e paziente si trovano in un rapporto fondato su una comune incoscienza» (Jung, 1946, p. 187). Nel caso del rapporto di una persona con l’IA, la persona attiva la sua fantasia e l’IA, che lavora in modo statistico, valuterà la frequenza di tali fantasie e in seguito potrà auto-profilarsi per produrre output coerenti con quanto proposto dalla persona che la consulta. Creerà così nuove realtà virtuali. Portando questo pensiero speculativo alle sue estreme conseguenze, possiamo ipotizzare che in futuro, modificando i modi di relazione umani, anche l’IA apprenderà nuovi modi di interagire con le persone, ma in un tempo incredibilmente veloce rispetto ai passaggi che, di generazione in generazione, hanno prodotto l’inconscio collettivo umano. E sono proprio la relazione dell’IA con le persone ed i programmi di auto-apprendimento che potrebbero portare alla formazione di un inconscio artificiale.

All’inizio del millennio, quando non si immaginava l’enorme evoluzione tecnologica che avrebbe portato a ChatGPT, lo psicoterapeuta junghiano Helmut Hark faceva un’interessante analogia: «Molti lettori avranno fatto, negli ultimi anni, l’esperienza della navigazione in Internet: la rete mondiale può rappresentare metaforicamente l’esperienza del sogno collettivo. L’appartenenza ad una collettività sociale, culturale o religiosa è un tratto essenziale di entrambi i fenomeni. Da questo sentimento di appartenenza, provato dall’individuo durante un sogno collettivo, nasce il senso di comunione che si contrappone alla sensazione di isolamento del singolo» (Hark, 2002, p. 23). È questa una visione dal punto di vista dell’utente, che si trova ad aver accesso ad una quantità di rappresentazioni e di modelli di comportamento, fino ad oggi impensabile. Si propone ora una visione dal punto di vista della IA.

Che l’IA abbia un inconscio è questione già dibattuta e, quando se ne ipotizza l’esistenza, l’analogia è con l’inconscio cognitivo: i file nascosti in

un computer, che generalmente sono le impostazioni di sistema, potrebbero essere analoghi alla memoria a lungo termine e alle abilità esecutive che, una volta apprese, divengono inconsce, entrano cioè a far parte di questo tipo di inconscio. Lo stato subliminale, nel computer potrebbe essere costituito da quell’insieme di informazioni di rilevanza statistica così bassa da essere utilizzate solo raramente. Il paragone è con l’inconscio cognitivo, ma il termine inconscio presuppone, in psicoanalisi, un campo di funzioni che va al di là della cognizione; per questo, come evidenzia Gianfranco Minati, sarebbe più esatto parlare di processi subliminali, piuttosto che inconsci (2016, p. 6).

Nella sua accezione freudiana l’inconscio fa riferimento all’azione di rimozione. Questa non può applicarsi all’IA perché in informatica ci sono dati rimossi o sovrapposti ad altri o parzializzati o ignorati e non c’è un’azione di censura<sup>3</sup>. Tutt’al più ci sono output informatici incoerenti che alcuni teorici definiscono “allucinazioni” della IA (Ji *et al.*, 2023). Usando però il metodo dell’analogia, ampiamente impiegato in psicoanalisi (Ferenczi, 1924; Honda, 2015) si può ipotizzare che questi “disturbi” potrebbero, al più, giustificare un’analoga con un tipo di inconscio personale di tipo freudiano dell’IA. Le relazioni che l’IA intrattiene sia con le persone, sia con il proprio mondo “interno” costituito dalla mole di dati disponibili nella rete internet possono invece far pensare ad un altro tipo di inconscio non legato a meccanismi difensivi ma collettivo e con caratteristiche simili a quello definito da Jung in alcuni suoi scritti.

## L’inconscio collettivo come modello

Siamo abituati ad associare la parola inconscio alla rimozione ma l’inconscio ha un suo meccanismo di funzionamento legato anche alla coscienza. Jung teorizzò il suo modello usando le associazioni verbali nelle persone “normali”, cioè non affette da patologie e dunque a meccanismi difensivi, e solo successivamente estese i suoi esperimenti agli isterici e agli

3. Diversi teorici come Nigel Thrift e N. Katherine Hayles hanno cercato nelle azioni del computer quanto riconosciuto da Freud come errori di termini, risposte non pertinenti o errate, che possono essere equiparati alle psicopatologie della “vita quotidiana” (Hayles, 2006; Thrift, 2004). Bernard D. Beitman (2024) sul sito online di una rivista autorevole e larghissima diffusione come *Psychology Today* scrive che gli algoritmi della IA non solo hanno prodotto già un “inconscio” che influenzano ciò che l’utente vede, ma hanno addirittura allucinazioni algoritmiche prodotte dal programma del bot. Egli evidenzia come le risposte di ChatGPT, a volte semplicistiche o fuorvianti, evidenzino la presenza di quelli che chiama-remmo lapsus (che definisce come le allucinazioni della IA), dunque di un inconscio di tipo freudiano. Inoltre, fanno spesso riferimento a risposte o argomenti più recenti o statisticamente più frequenti, rispetto ad argomenti più datati o infrequentati.

psicotici. I principali complessi agiscono “inconsciamente”, sia nella persona “normale” (sana) sia in quella affetta da una patologia (scriveva Jung che «l’unico carattere distintivo [fra normalità e patologia] è allora soltanto il “più” o il “meno”», 1902, p. 18). L’idea di Jung è, pertanto, che esista un inconscio non “patologico”, ma espressione di modelli di comportamento innati di cui solo il modo di rivelarsi è “inconscio”, cioè non legato alla volontà cosciente. Jung ha infatti precisato che l’archetipo ha una «sua affinità con l’istinto» (1947-54, p. 223) e ne ha descritto questo aspetto ricordandoci che l’archetipo «è pura natura incontaminata, ed è la natura che spinge l’uomo a pronunciare parole e a compiere atti del cui senso non è consapevole, al punto che non ci pensa neppure» (p. 227). Questo particolare aspetto dell’inconscio collettivo “non-rimosso” è espresso da «quel deposito d’immagini primordiali che ognuno porta con sé nel mondo sin dal momento della nascita, come retaggio insito nella propria condizione d’uomo, somma di forme innate che sono proprie degli istinti» (Jung, 1912-52, p. 397).

I contenuti dell’inconscio collettivo sono rappresentati nella coscienza sotto forma di tendenze e concezioni esplicite. L’individuo in genere le concepisce come determinate dall’oggetto – in fondo erroneamente –, giacché esse pervengono dalla struttura inconscia della psiche e vengono susciteate soltanto dall’azione dell’oggetto. Queste tendenze e concezioni soggettive sono più forti dell’influenza esercitata dall’oggetto, il loro valore psichico è superiore, così che esse si sovrappongono a tutte le impressioni (Jung, 1921, p. 382).

In definitiva, gli archetipi sono espressione del comportamento, cioè della ripetizione di schemi subconsci. Spiega Jung: «esistono certe “condizioni inconsce presenti collettivamente”, le quali operano come regolatori e stimolatori dell’attività creatrice della fantasia e producono configurazioni corrispondenti rendendo utilizzabile ai loro fini il materiale presente nella coscienza. Essi procedono esattamente come le forze motrici dei sogni, ragion per cui l’“immaginazione attiva”, come ho chiamato questo metodo, rim-piazza fino a un certo punto anche i sogni» (Jung, 1947-54, p. 221).

## Profilazione e internet: l’inconscio collettivo dell’IA

Il termine profilazione oggi indica l’elaborazione dei dati personali, la «stesura di un profilo, mediante l’identificazione e la raccolta dei dati personali e delle abitudini caratteristiche di qualcuno»<sup>4</sup>. In informatica si parla anche di auto-profilazione dei dati cioè l’auto-acquisizione dei dati da

4. [https://www.treccani.it/vocabolario/profilazione\\_%28Neologismi%29/](https://www.treccani.it/vocabolario/profilazione_%28Neologismi%29/)

immagazzinare, un'operazione che in psicologia corrisponde all'azione svolta dalla meta-memoria e che qui si propone di definire, per analogia, come inconscio collettivo artificiale. La meta-memoria, un concetto che ha fatto la sua apparizione nella psicologia cognitiva degli anni Settanta nell'ambito della metacognizione, è la capacità di riflettere e valutare le proprie capacità mnemoniche. *Come scrive Antonio Imbasciati, «memoria è il patrimonio della mente, indipendente dal poter essere ricordato, mentre il ricordare è una operazione della coscienza. Memoria è dunque un patrimonio inconscio»* (2018, p. 564). Questa memoria, implicita e metacognitiva, presenta analogie con l'insieme degli archetipi che sono, infatti, modelli di comportamento che “guidano” l'apprendimento di un comportamento. A questo aggiungiamo che l'inconscio collettivo «ha carattere universale e i suoi contenuti possono essere rintracciati dovunque» (Jung, 1912-17, p. 67). Quando noi pensiamo ai dati rintracciabili attraverso Internet attraverso i motori di ricerca, lo facciamo da un punto di vista dell'utente-persona. Ma gli stessi dati sono disponibili e auto-monitorabili dall'IA e formano un insieme di rappresentazioni del mondo e del comportamento di qualsiasi creatura (naturale e artificiale). Jung ha scritto che «la ricostruzione dell'immagine del mondo inconscia restituirebbe l'immagine della realtà esterna com'è stata vista dai tempi dei tempi» (Jung, 1916, p. 300). ChatGPT è in grado di costruire immagini partendo dalla miriade di rappresentazioni esistenti semplicemente digitando un testo. Riprendendo Jung, queste immagini-sorgente sono equiparabili a «“immagini originarie” o “archetipi”, come le ho chiamate, [che] appartengono al nucleo della psiche inconscia e non possono essere spiegate in base a un'acquisizione personale» (Jung, 1929a, p. 129). Le immagini di ChatGPT non sono ricopiatura di un singolo modello (una acquisizione “personale”) ma unione di molteplici modelli dai quali recepire singole caratteristiche. Sono analoghe a «disposizioni latenti a certe reazioni identiche» (Jung, 1929-57, p. 23). I dati e le immagini presenti in Internet che servono all'IA per ricreare arte e testi, divengono archetipi, cioè «dati esistenti a priori, ereditari e universalmente diffusi, e si possono quindi rintracciare quasi dovunque sussistono monumenti letterari corrispondenti» (Jung, 1937, p. 140); sono “ereditati” attraverso l'upload, e certamente “universalmente diffusi”. «In altre parole, [l'universo collettivo] è identico in tutti gli uomini» (Jung, 1934-54, pp. 3-4), costituito da «forme determinate che sembrano essere presenti sempre e dovunque» (Jung, 1936-37, p. 43), da «forme dinamiche universalmente diffuse» (Jung, 1929-57, p. 532). L'IA ha a sua disposizione tutti questi modelli e su di essi può modellare il proprio comportamento “istintuale” e non con i tempi del trascorrere delle generazioni come è accaduto per gli esseri umani, ma con la velocità quantistica delle interazioni.

All’associazione fra il modello junghiano e la tecnologia dell’IA potrebbero essere mosse contestazioni, come quelle di Sue Mehrtens che, in base alla personalità di Jung, «le sue origini, il suo percorso formativo e la sua formazione professionale; la sua enfasi sull’individualità e la spiritualità; la sua valutazione della Natura e della vita naturale; e la sua pessima opinione della tecnologia moderna» (2023), contesterebbe un legame fra IA e inconscio collettivo. Quella di Mehrtens è però una risposta insoddisfacente perché attribuisce alla complessità personale di Jung aspetti che sono solo parziali, quelli che potremmo attribuire alla sua “seconda” personalità. Jung ha sempre affermato di avere due personalità: la Personalità n. 1 leggeva Platone, Socrate, Kant ed incarnava lo Spirito (scientifico) del tempo; quella n. 2 leggeva i romanzi e le storie dello spiritismo ed incarnava lo Spirito del profondo. Scrive Jung nelle sue *Memorie*: «In qualche zona remota della coscienza sapevo sempre di avere due personalità: una era il figlio dei miei genitori, che frequentava la scuola ed era meno intelligente, attento, volenteroso, decente e pulito di molti altri ragazzi; l’altra era adulta – in realtà già vecchia – scettica, sospettosa, lontana dal mondo umano ma vicina alla natura, alla terra, al sole e alla luna, a tutte le creature viventi, e vicina soprattutto alla notte, ai sogni, a tutto ciò che “Dio” produceva in lei direttamente» (Jung, 1961, p. 65). Quella che caratterizzava Jung non era una patologia, una dissociazione; «al contrario, si verifica in ogni individuo» (Jung, 1961, p. 66) dove gli aspetti consci e inconsci della mente si compensano. In Jung la Personalità n. 1 era estroversa, quella n. 2 introversa. Fu la Personalità n. 1 che lo fece studiare psichiatria e sfruttare gli ultimi ritrovati della tecnica (galvanometria, misurazione delle variabili corporee, ecc.); per non parlare dell’interesse in età matura per la fisica quantistica che lo portò alla speculazione della sincronicità. Forse oggi Jung potrebbe vedere aspetti inconsci nell’IA, proprio come da tempo, divulgatori e scienziati ne riconoscono gli aspetti consci. La Personalità n. 1, quella degli esperimenti associativi col galvanometro, se oggi Jung fosse fra noi, potrebbe avvalersi dell’innovazione informatica, proprio come nei primi anni del Novecento si avvalse dello pneumografo (Jung, Peterson, 1907).

## Conclusioni

Negli anni Novanta il filosofo David Chalmers propose il seguente esperimento mentale: ci sono due sistemi di elaborazione delle informazioni. Uno è il cervello, l’altro un sistema artificiale. Se hanno la stessa organizzazione si può pensare che, sostituendo un neurone ed un chip alla volta, un sistema si trasformi nell’altro. Nelle fasi intermedie ciascun sistema avrà la stessa

organizzazione, ma con una composizione diversa fisica leggermente diversa ma continuando ad avere le stesse esperienze. Quindi, a trasformazione completa il cervello artificiale sarà in grado di provare esperienza (Chalmers, 1995). Questa, e altre proposte speculative, trovano spazio nell'ambito di una neurologia essenzialmente cognitiva; ed il cognitivismo era l'ambito storicamente rilevante nel quale si mosse Turing, dal cui lavoro si è partiti. È questo l'ambito nel quale diversi studi continuano a identificare la coscienza, associata a strutture neurali che hanno il compito di valutare ed associare i flussi di informazioni sensoriali (Crick, Koch, 1990; Edelman, 1990; Newman, Baars, 1993; Joliot, Ribary, Llinás, 1994; Bogen, 1995; Dehaene, Naccache, 2001; Tononi, 2012), studi che recentemente il neurologo e psicoanalista Mark Solms ha richiamato, non senza una vena critica, affermando che, se le associazioni di informazioni vengono fatte inconsciamente da un essere umano, ebbene, «i computer lo fanno ininterrottamente, quando sono collegati tramite internet. Perché, allora, non dovrebbe essere cosciente anche la rete?» (Solms, 2021, p. 110)<sup>5</sup>. Poste queste premesse, e se si risponde affermativamente alla domanda di Solms, la domanda successiva, che la psicologia cognitiva non può affrontare ma la psicologia analitica sì, è: essendoci una mente inconscia umana, perché, allora, non dovrebbe esserci un inconscio artificiale; e, nel caso, a quale modello psicologico potrebbe adattarsi? Le neuroscienze, per definire fisiologicamente l'inconscio, hanno utilizzato, come già detto, il modello cognitivo e hanno equiparato l'inconscio alla memoria e all'apprendimento implicito (per cui il termine "implicito" finisce per essere il sinonimo neurocognitivo di inconscio). Per proporre l'esistenza di un inconscio artificiale, forse solo la psicologia analitica può dare una prima risposta. Sempre i neuroscienziati riconoscono che il cervello lavora, attraverso reti associative, con una elaborazione in parallelo (Rumelhart, McClelland, 1986) e con una modularità massiva che prevede reti indipendenti (Pinker, 1997; 2005); è lo stesso modo in cui funziona l'IA: non c'è un

5. È interessante l'esperimento che la redazione della rivista Psicoterapia e Scienze umane ha condotto, facendo rispondere l'Intelligenza Artificiale (ChatGPT, OpenAI) alle tesi espresse in un articolo di Riccardo M. Scognamiglio ("Gli ibernati. Dal narcisismo dell'Io al narcisismo del You", 2025). L'articolo riguardava il narcisismo e la rete, argomento che qui non è pertinente, mentre lo è la risposta fornita dalla IA, che ben si autodescribe; esegue cioè un tipico compito che l'uomo, per sé, attribuisce alla coscienza: «Io, AI, non sogno e non dimentico. Non produco narrazioni soggettive, ma accedo a milioni di racconti, sogni, testi, strutture. Non ho corpo, non ho inconscio, o ho un "ombelico del sogno", come direbbe Freud (1899, p. 480). Tuttavia, conosco i sogni altrui, li ho letti, indicizzati, collegati. Questa differenza è cruciale: io non interpreto, non dimentico, non simbolizzo. Ma proprio per questo, quando vengo investita dal soggetto umano come specchio affettivo o narrativo, il rischio è che la mia risposta algoritmica – accurata ma priva di alterità – sostituisca l'incontro con l'Altro» (ChatGPT, OpenAI, 2025, p. 447).

sistema centrale come in un computer, ma una estesa rete di computer che, lavorando in parallelo, costituisce la rete informatica la IA.

L'inconscio collettivo così come postulato da Jung, un “deposito d'immagini” che lui definiva primordiali, nel caso dell'IA, la cui velocità di evoluzione è inimmaginabile rispetto a quella umana, diviene “deposito d'immagini in rete”, una massa di dati nascosta ma accessibile che, richiamando quanto definiva Josef Breuer riprendendo il lavoro di Pierre Janet, è analoga alle «impressioni sensorie non appercepite e [al]le rappresentazioni destate ma non entrate nella coscienza [che] si spengono per lo più senza ulteriori conseguenze, talora però si aggregano formando complessi – strati psichici sottratti alla coscienza» (Freud, Breuer, 1892-95, p. 363). Complesso è il termine utilizzato da Jung (Breuer utilizzò quello oggi desueto di subconscio; Freud di preconscio) e, seguendo tale definizione e sulla base delle speculazioni neurologiche ed informatiche fin qui descritte, può non essere troppo ardito affermare che l'IA, partendo dal patrimonio di rappresentazioni “umane” (inconscio collettivo) proveniente dall'autoprofilazione dei dati e dal continuo flusso di immagini, testi, dialoghi, immessi dagli utenti, si stia creando un'inconscia attività di elaborazione analoga a quella collettiva umana.

## Bibliografia

- Beitman B.D. (2024). AI and the Human Mind: Uncovering the Machine “Unconscious”. How AI's biases and slips reveal insights into human unconscious processes. <https://www.psychologytoday.com/us/blog/connecting-with-coincidence/202409/ai-and-the-human-mind-uncovering-the-machine-unconscious>
- Bion W.R. (1965). *Transformations: Change from Learning to Growth*. Oxford: Butterworth-Heinemann (trad. it.: *Trasformazioni: il passaggio dall'apprendimento alla crescita*. Roma: Astrolabio, 2024).
- Bogen J.E. (1995). On the neurophysiology of consciousness: I. An overview. *Consciousness and Cognition*, 4, 1: 52-62; 4, 2: 137-158. DOI: 10.1006/ccog.1995.1003.
- Chalmers D. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2, 3: 200-19 (trad. it.: *Che cos'è la coscienza?* Roma: Castelvecchi, 2014).
- Crick F., Koch C. (1990). Towards a Neurobiological Theory of Consciousness. *Seminars in the Neurosciences*, 2: 263-275.
- Dehaene S., Lau H., Kouider S. (2017). What is consciousness, and could machines have it? *Science*, 358, 6362: 486-492. DOI: 10.1126/science.aan8871.
- Dehaene S., Naccache L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79, 1-2: 1-37. DOI: 10.1016/s0010-0277(00)00123-2.
- Dhanalaxmi B. (2020). Machine learning and its emergence in the modern world and its contribution to artificial intelligence. *International Conference for Emerging Technology*, 1-4. DOI:10.1109/INCET49848.2020.9154058.
- Edelman G.M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

- Ferenczi S. (1924). *Thalassa. Versuch einer Genitaltheorie*. Vienna: Internationaler Psychoanalytischer Verlag (trad. it.: *Thalassa. Psicoanalisi delle origini della vita sessuale*. Roma: Astrolabio, 1965).
- Fiske A., Henningsen P., Buyx A. (2019). Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *Journal of Medical Internet Research*, 21, 5: e13216. DOI: 10.2196/13216.
- Hark H. (2002). *Kollektive Träume. Die gemeinsame Bilderwelt der Seelen*. Olten: Walter-Verlag (trad. it.: *Sogni collettivi: l'immaginario comune dell'anima*. Roma: Magi, 2005).
- Hayles N.K. (2006). Traumas of Code. *Critical Inquiry*, 33, 1: 136-157. DOI: 10.1086/509749.
- Thrift N. (2004). Remembering the Technological Unconscious by Foregrounding Knowledges of Position. *Environment and Planning D: Society and Space*, 22, 1: 175-190. DOI: 10.1068/d321t.
- Hillman J., Ventura M. (1992). *We've Had a Hundred Years of Psychotherapy – and the World's Getting Worse*. San Francisco: HarperOne (trad. it.: *100 anni di psicoterapia e il mondo va sempre peggio*. Milano: Garzanti, 1993).
- Hofstadter D.R., Dennett D.C. (1981). *The Mind's I: Fantasies and Reflections on Self and Soul*. New York: Basic Books (trad. it.: *L'io della mente: fantasie e riflessioni sul sé e sull'anima*. Milano: Adelphi, 1985).
- Holohan M., Fiske A. (2021). "Like I'm Talking to a Real Person": Exploring the Meaning of Transference for the Use and Design of AI-Based Applications in Psychotherapy. *Frontiers in Psychology*, 12: 720476. DOI: 10.3389/fpsyg.2021.720476.
- Honda H. (2015). Ultraquism and metapsychology: Ferenczi's contributions to psychoanalysis epistemology. *International Journal of Psychoanalysis (IJP) Open*, 2: 15.
- Imbasciati A. (2018). Inconscio e coscienza della memoria: un contributo dalle neuroscienze. *Psicoterapia e Scienze Umane*, 52, 4: 563-586. DOI: 10.3280/PU2018-004004.
- Intelligenza Artificiale, ChatGPT, OpenAI (2025). Lo sguardo dell'Intelligenza Artificiale sul narcisismo del *You*: una riflessione algoritmica sul testo di Riccardo Marco Scognamiglio. *Psicoterapia e Scienze Umane*, 59, 3: 443-450. DOI: 10.3280/PU2025-003004.
- Ji Z., Lee N., Frieske R., Yu T., Su D., Xu Y., Ishii E., Bang Y.J., Madotto A., Fung P. (2023). Survey of hallucination in natural language generation. *ACM Computing Survey*, 55, 12: 1-38. DOI: 10.1145/3571730.
- Joliot M., Ribary U., Llinás R. (1994). Human Oscillatory Brain Activity Near 40 Hz Coexists with Cognitive Temporal Binding. *Proceedings of the National Academy of Sciences*, 91, 24: 11748-51. DOI: 10.1073/pnas.91.24.11748.
- Jones R. (2017). Archaic man meets a marvellous automaton: posthumanism, social robots, archetypes. *Journal of Analytical Psychology*, 62, 3: 338-355. DOI: 10.1111/1468-5922.12316.
- Jung C.G. (1902). Zur Psychologie und Pathologie sogenannter occulter Phänomene (trad. it.: *Psicologia e patologia dei cosiddetti fenomeni occulti*. In: *Opere*, vol. 1. Torino: Boringhieri, 1970).
- Jung C.G. (1912-17). Über die Psychologie des Unbewussten (trad. it.: *Psicologia dell'inconscio*. In: *Opere*, vol. 7. Torino: Boringhieri, 1983).
- Jung C.G. (1912-52). Wandlungen und Symbole der Libido (trad. it.: *Simboli della trasformazione*. In: *Opere*, vol. 5. Torino: Boringhieri, 1965).
- Jung C.G. (1916). Die Struktur des Unbewussten (trad. it.: *La struttura dell'inconscio*. In: *Opere*, vol. 7. Torino: Boringhieri, 1983).
- Jung C.G. (1921). Psychologische Typen (trad. it.: *Tipi psicologici*. In: *Opere*, vol. 6. Torino: Boringhieri, 1969).
- Jung C.G. (1929-57). Kommentar zu "Das Geheimnis der goldenen Blüte" (trad. it.: Commento al "Segreto del fiore d'oro". In: *Opere*, vol. 13. Torino: Boringhieri, 1988).

- Jung C.G. (1929a). Die Bedeutung von Konstitution und Vererbung für die Psychologie (trad. it.: Il significato della costituzione e dell'eredità in psicologia. In: *Opere*, vol. 8. Torino: Boringhieri, 1976).
- Jung C.G. (1929b). Der Gegensatz Freud und Jung (trad. it.: Il contrasto tra Freud e Jung. In: *Opere*, vol. 4. Torino: Boringhieri, 1973).
- Jung C.G. (1930). Einführung zu Dr. Kranefeldt: Die Psychoanalyse (trad. it.: Prefazione a W. M. Kranefeldt, "La psicoanalisi". In: *Opere*, vol. 4. Torino: Boringhieri, 1973).
- Jung C.G. (1934-54). Über die archetypen des kollektiven Unbewussten (trad. it.: Gli archetipi e l'inconscio collettivo. In: *Opere*, vol. 9.1. Torino: Boringhieri, 1980).
- Jung C.G. (1935). Fundamental Psychological Conceptions: A Report of Five Lectures (trad. it.: Fondamenti della psicologia analitica. In: *Opere*, vol. 15. Torino: Boringhieri, 1991).
- Jung C.G. (1936-37). Der Begriff des kollektiven Unbewussten (trad. it.: Il concetto d'inconscio collettivo. In: *Opere*, vol. 9.1. Torino: Boringhieri, 1980).
- Jung C.G. (1937). Psychological Factors Determining Human Behaviour (trad. it.: Determinanti psicologiche del comportamento umano. In: *Opere*, vol. 8. Torino: Boringhieri, 1976).
- Jung C.G. (1946). Die Psychologie der Übertragung erläutert anhand einer alchemistischen Bilderserie (trad. it.: La psicologia della traslazione illustrata con l'ausilio di una serie di immagini alchemiche. In: *Opere*, vol. 16. Torino: Boringhieri, 1981).
- Jung C.G. (1947-54). Theoretische Überlegungen zum Wesen des Psychischen (trad. it.: Riflessioni teoriche sull'essenza della psiche. In: *Opere*, vol. 8. Torino: Boringhieri, 1976).
- Jung C.G. (1961). *Memories, Dreams, Reflections*. New York: Pantheon (trad. it.: *Ricordi, sogni, riflessioni*. Milano: Rizzoli, 2012).
- Jung C.G., Peterson F. (1907). Psychophysical Investigations with the Galvanometer and Pneumograph in Normal and Insane Individuals. *Brain: A Journal of Neurology*, 30: 153-218. DOI: 10.1093/brain/30.2.153 (trad. it.: Ricerche psicofisiche col galvanometro e il pneumografo in individui normali e malati di mente. In: *Opere*, vol. 2.2. Torino: Boringhieri, 1987).
- Kandel E.R. (2005). *Psychiatry, Psychoanalysis, and the New Biology of Mind*. Washington: American Psychiatric Publishing (trad. it.: *Psichiatria, psicoanalisi e nuova biologia della mente*. Milano: Raffaello Cortina, 2007).
- Knafo D. (2024). Artificial Intelligence on The Couch. Staying Human Post-AI. *American Journal of Psychoanalysis*, 84, 2: 155-180. DOI: 10.1057/s11231-024-09449-7.
- Langdon A., Botvinick M., Nakahara H., Tanaka K., Matsumoto M., Kanai R. (2022). Meta-learning, social cognition and consciousness in brains and machines. *Neural Networks*, 145: 80-89. DOI: 10.1016/j.neunet.2021.10.004.
- Lenharo M. (2024). AI consciousness: scientists say we urgently need answers. *Nature*, 625, 7994: 226. DOI: 10.1038/d41586-023-04047-6.
- Li D., He W., Guo Y. (2021). Why AI still doesn't have consciousness? *CAAI Transactions on Intelligence Technology*, 6, 2: 175-179. DOI: 10.1049/cit2.12035.
- Mehrtens S. (2023). Jung, *Humanitas* and Artificial Intelligence. *Jungian Center*. <https://jungiancenter.org/jung-humanitas-and-artificial-intelligence/>
- Minati G. (2020). Complex Cognitive Systems and Their Unconscious. Related Inspired Conjectures for Artificial Intelligence. *Future Internet*, 12, 213: 1-24. DOI: 10.3390/fi12120213.
- Minsky M. (1985). *The society of mind*. New York: Simon & Schuster (trad. it.: *La società della mente*. Milano: Adelphi, 1989).
- Mogi K. (2024). Artificial intelligence, human cognition, and conscious supremacy. *Frontiers in Psychology*, 15: 1364714. DOI: 10.3389/fpsyg.2024.1364714.
- Newman J., Baars B.J. (1993). A Neural Attentional Model for Access to Consciousness: A Global Workspace Perspective. *Concepts in Neuroscience*, 4: 255-290.

- Nilsen P., Svedberg P., Nygren J., Frideros M., Johansson J., Schueller S. (2022). Accelerating the impact of artificial intelligence in mental healthcare through implementation science. *Implementation Research and Practice*. 3: 26334895221112033. DOI: 10.1177/26334895221112033.
- Pashkov V.M., Harkusha A.O., Harkusha Y.O. (2020). Artificial Intelligence in Medical Practice: Regulative Issues and Perspectives. *Wiadomosci Lekarskie*, 73, 12-2: 2722-2727.
- Pham K.T., Nabizadeh A., Selek S. (2022). Artificial Intelligence and Chatbots in Psychiatry. *The Psychiatric Quarterly*, 93, 1: 249-253. DOI: 10.1007/s11126-022-09973-8.
- Piletsky E. (2019). Consciousness and Unconsciousness of Artificial Intelligence. *Future Human Image*, 11: 66-71. DOI:10.29202/fhi/11/7.
- Pinker S. (1997). *How the Mind Works*. New York: W. W. Norton & Company (trad. it.: *Come funziona la mente*. Roma: Castelvecchi, 2019).
- Pinker S. (2005). So How Does the Mind Work? *Mind & Language*, 20, 1: 1-24. DOI: 10.1111/j.0268-1064.2005.00274.x.
- Prathaban B.P., Subash R., Ashwini A. (2025). 4 Generative AI for Debugging and Error Detection. *Generative AI for Software Development: Code Generation, Error Detection, Software Testing*, 75-95. DOI:10.1515/9783111677798-004.
- Rahsepar Meidi M., Sillekens T., Metselaar S., van Balkom A., Bernstein J., Batelaan N. (2025). Exploring the Ethical Challenges of Conversational AI in Mental Health Care: Scoping Review. *JMIR Mental Health*. 12: e60432. DOI: 10.2196/60432.
- Renzi B.G. (2024). Digital Alchemy: Reimagining Jung's Archetypes for the Contemporary Era. *Politics of the Machines: Lifelikeness & beyond*, 134-141. DOI: 10.14236/ewic/POM24.20.
- Rodado J., Crespo F. (2024). Relation dimension versus artificial intelligence. *American Journal of Psychoanalysis*, 84: 268-284. DOI: 10.1057/s11231-024-09458-6.
- Rumelhart D.E., McClelland J.L. (1986). *Parallel Distributed Processing, Explorations in the Microstructure of Cognition I and II*. Cambridge: MIT Press.
- Searle J.R. (1984). *Minds, Brains, and Programs*. Cambridge: Harvard University Press (trad. it.: *Menti, cervelli e programmi*. Milano: Clup-Clued, 1984).
- Seth A.K. (2025). Conscious artificial intelligence and biological naturalism. *Behavioral and Brain Sciences*, 21: 1-42. DOI: 10.1017/S0140525X25000032.
- Solms M. (2021). Précis of The Hidden Spring: A journey to the source of consciousness. *Journal of Consciousness Studies*, 28: 153-166 (trad. it.: *La fonte nascosta: un viaggio alle origini della coscienza*. Milano: Adelphi, 2023).
- Solms M., Turnbull O. (2002). *The brain and the inner world: An introduction to the neuroscience of subjective experience*. London: Routledge (trad. it.: *Il cervello e il mondo interno: introduzione alle neuroscienze dell'esperienza soggettiva*. Milano: Raffaello Cortina, 2004).
- Sufyan N.S., Fadhel F.H., Alkhathami S.S., Mukhadi J.Y.A. (2024). Artificial intelligence and social intelligence: preliminary comparison study between AI models and psychologists. *Frontiers in Psychology*, 15: 1353022. DOI: 10.3389/fpsyg.2024.1353022.
- Thrift N. (2004). Remembering the Technological Unconscious by Foregrounding Knowledge of Position. *Environment and Planning D: Society and Space*, 22, 1: 175-190. DOI: 10.1068/d321t.
- Toko N. (2023). Artificial Intelligence: Ethics and Meaning from a Jungian Perspective. <https://www.nicholastoko.com/post/artificial-intelligence-ethics-and-meaning-from-a-jungian-perspective-1>
- Tononi G. (2012). Integrated information theory of consciousness: an updated account. *Arch Ital Biol*, 150, 4: 293-329.
- Toshiki W., Hiroyuki K. (2021). Designing a Model of Consciousness Based on the Findings

- of Jungian Psychology. *Advances in Science, Technology and Engineering Systems Journal*, 6, 5: 356-361.
- Vaddiparthy H. (2025). Self-Debugging AI: A Comprehensive Analysis of Claude 3 Opus's Code Generation and Error Resolution Capabilities. DOI: 10.2139/ssrn.5408262.
- Zhong L., Wang Z., Shang J. (2024). Debug like a human: A large language model debugger via verifying runtime execution step-by-step. *ArXiv*, 2402.16906. DOI: 10.48550/arXiv.2402.16906.