

Pastiche visuali e intelligenza artificiale. Iporealismo, pittoresco, simulacri

Daniele Villa

Politecnico di Milano, Dipartimento di Architettura e Studi Urbani
(daniele.villa@polimi.it)

In un lasso di tempo molto breve, sul finire del 2022, le ormai vetuste relazioni fra noi e le macchine sono nuovamente e profondamente mutate: nei mesi di gennaio e febbraio 2023 oltre cento milioni di persone hanno iniziato a familiarizzare con il primo e più diffuso *generatore di testo preistruito* basato su una serie di algoritmi di auto-apprendimento e intelligenza artificiale: ChatGPT, della società californiana OpenAI.

Si pone al sistema una domanda qualsiasi e ChatGPT risponde dalla sua interfaccia web con una naturalezza, un eloquio, a volte una esautività sconcertanti. Nel prendere atto di una serie di nuove e ampiamente inesplorate potenzialità dell'intelligenza artificiale possiamo scegliere se rimanere estasiati dalla retorica dominante della nuova e progressiva irrefrenabile potenza dell'artificio informatico, o fare lo sforzo di coltivare la scomodità di un parziale disincanto, cercando nelle pieghe della meraviglia alcuni spunti di riflessione. Gli algoritmi di *machine-learning*, da un decennio almeno fra le più promettenti innovazioni del mondo digitale, sono improvvisamente passati di mano: non più appannaggio di una cerchia di esperti, tecnici e nerd entusiasti, ma liberamente fruibili dalla generalità di ognuno di noi, apparentemente senza limiti nelle possibilità di sperimentazione e ideazione.

Sistemi automatici che *imparano* da noi, rispondono alle nostre domande, ci trattano con affettata gentilezza, compongono testi, poesie, righe di codice informatico; interagiscono, si correggono, riassumono o ampliano concetti, generano *ex nihilo* suoni, musiche, immagini. Come fa notare non senza un certo sarcasmo Noam Chomsky in un recente e implacabile articolo apparso sul *New York Times*: «OpenAI's ChatGPT, Google's Bard and Microsoft's Sydney are marvels of machine learning. Roughly speaking, they take huge amounts of data, search for patterns in it and become increasingly proficient at generating statistically probable outputs — such as seemingly humanlike language and thought. These programs have been hailed as the first glimmers on the horizon of artificial general intelligence — that long-prophesied moment when mechanical minds surpass human brains not only quantitatively in terms of processing speed and memory size but also qualitatively in terms of intellectual insight, artistic creativity and every other distinctively human faculty» (Chomsky, Roberts, Watmull, 2023).

Siamo di fronte alla nascita di quale *intelligenza*? In che cosa, significativamente, cambierà il nostro approccio alla conoscenza attraverso questi strumenti, sempre più pervasivi, di informazione automatica aumentata? Proviamo qui a ipotizzare alcune risposte

da un punto di vista apparentemente eccentrico, prendendo spunto dalle connessioni che legano la spazializzazione del nostro conoscere (nel territorio, nella città, nel paesaggio, nell'architettura) e le più diverse accezioni della sua codificazione visuale. Questo scorcio non è casuale, ma muove da un presupposto empirico: i due principali canali di fruizione generalista degli algoritmi di intelligenza artificiale sono a oggi rappresentati dall'interazione umano-macchina attraverso la generazione automatica di testo e, parallelamente, tramite modelli generativi che dal testo non producono altro testo, ma immagini. Decine di milioni di immagini prodotte quotidianamente attraverso svariati e crescenti servizi concorrenti di *text-to image artificial intelligence*.

Tutto inizia, e finisce, con la crescita esponenziale della capacità di calcolo nell'elaborare dati, enormi quantità di dati, senza i quali gli algoritmi non potrebbero imparare a riprodurre e catalogare le caratteristiche visive di miliardi di fonti, poi scomposte e ricomposte in qualcosa che assomigli, sempre di più e sempre meglio, a un significato il cui significante artificiale è soggetto a ulteriori infiniti rimescolamenti digitali (cfr. Manovich, 2023). La sintesi eidomatica di immagini non è affatto una novità, ma operazioni di *mixing*, *morphing* o arte generativa che solo pochi anni fa richiedevano giorni di lavoro, specifiche competenze tecniche e macchinari molto costosi sono ora realizzabili sostanzialmente da chiunque, in pochi secondi. A complicare la faccenda c'è poi il cortocircuito semantico prodotto da modelli di generazione di immagini basati su input di testo, nei quali i possibili risultati mutano totalmente cambiando un solo aggettivo o arricchendo un dettaglio descrittivo. Si tratta di capire come andare oltre la legittima fascinazione collettiva data da questa accelerazione improvvisa delle possibilità offerte dall'intelligenza artificiale cercando un campo di possibili nuove opportunità e nuovi interrogativi che, come ci ricorda Tomás Maldonado «travalicano ampiamente i confini della tecnologia» (Maldonado, 1992: 8). Il medium visuale risulta in questo caso dicotomicamente adatto tanto a simulare grandi innovazioni, quanto a permetterci di individuare sottotraccia meccanismi culturalmente molto radicati nella storia moderna dell'iconicità e delle sue derive. È di fatto banale constatare, a titolo di esempio, quanto le immagini create dai sistemi di *machine learning* non contengano alcuna effettiva novità geometrico-visuale: più del 99% di questi prodotti digitali ha una matrice prospettica e una struttura filmico-fotografica iper-consolidata e rassicurante, arricchita da un prodigioso livello di dettaglio realistico e dalla capacità di ibridazione quasi perfetta fra elementi incongruenti.

Tutto qui tende ad assomigliare moltissimo a qualcosa che abbiamo già visto, in una sorta di clonazione eterozigote illimitata e, proprio come accadde agli albori del pensiero prospettico: «Non importa tanto sapere, in astratto, se vi è una corrispondenza biunivoca totale tra una rappresentazione prospettica e la realtà che essa intende rappresentare. Più ragionevole è cercar di stabilire se la rappresentazione prospettica, per usare una parola cara all'operazionismo, funziona (o meno) come rappresentazione plausibile, non di una generica realtà, ma piuttosto della nostra percezione della realtà» (Maldonado, 1992:19).

In questa plausibile riproduzione autonoma della nostra percezione della realtà è difficile negare il ruolo centrale, convincente e mimetico, che la rappresentazione dello spazio è in grado di giocare. La stragrande maggioranza delle visioni prodotte dagli algoritmi generalisti di *machine learning* ha, di fatto, riferimenti allo spazio domestico, architettonico, urbano o paesaggistico. Si è quindi provato, con le quarantuno immagini che seguono, a sperimentare una forzatura ludica che obbligasse l'intelligenza artificiale a raccontarci il suo modo di *vedere* lo spazio, la città e il territorio, a prescindere da ogni altro contenuto, nel tentativo di individuare, sotto la superficie, possibili ricorrenze, elementi di interesse, specificità visuali o culturali degne di interesse.

La prima considerazione, a valle di questo esperimento, riguarda propriamente il tema della verosimiglianza. Come i modelli di apprendimento linguistico automatico mescolano sempre più efficacemente sequenze di parole a partire da un database testuale potenzialmente illimitato, alla ricerca di quelle forme lessicali che hanno la maggior probabilità di assomigliare a ciò che noi umani siamo abituati a dire e ad ascoltare, così i generatori automatici di immagini campionano in manciate di secondi milioni di fonti visive e ne fondono elementi scartando ciò che meno si adattano a una plausibile verosimiglianza. Ogni volto, ogni luogo, spazio, prodotto dalla AI è un *pastiche* così profondamente falso eppure così dettagliato e raffinato, nel senso letterale del termine, da apparire del tutto nuovo pur essendo incredibilmente somigliante a qualcosa che abbiamo già visto. Si innesca in questo modo un meccanismo parzialmente inatteso: non ci troviamo di fronte, come potrebbe sembrare, alla vertigine dell'iperrealismo, ma a un *déjà vu* ciclico, in cui tutto il *mai visto prima* non fa che avvicinarsi alla possibilità statistica di assomigliare a qualcosa di perfettamente noto.

Tanto conosciuti, alcuni di questi *neo-non-luoghi* artificiali da darci la sensazione di ricordarli, di esserci passati. Tocchiamo con mano, anzi con gli occhi, uno degli effetti collaterali più inquietanti prodotti dal *deep fake* elettronico, il cosiddetto Mandela Effect. Come scrive Andrea Indiano su Wired: «L'arte generata dall'intelligenza artificiale può così creare l'illusione dei ricordi, in particolare se è abbastanza realistica da essere scambiata per qualcosa che è accaduto veramente [...]. Questo perché il cervello umano è altamente suscettibile ai segnali visivi e può essere facilmente ingannato da immagini realistiche. I bot rischiano così di causare un Mandela Effect in chi osserva i loro lavori. Si tratta di un fenomeno in cui un gruppo di persone ricorda collettivamente un evento o un fatto in modo diverso da come è realmente accaduto. Il termine è stato coniato dalla ricercatrice Fiona Broome dopo aver scoperto che un numero significativo di persone ricordava che Nelson

Mandela fosse morto in prigione negli anni '80, quando in realtà fu rilasciato nel 1990 e morì nel 2013. Da allora la frase sta a indicare un ricordo collettivo fallace ed è stata spesso associata alla cultura pop» (Indiano, 2022).

Non è la realtà a essere imitata in modo tanto efficace, bensì il nostro stesso modo di guardarla, di elaborarla, di condividerla, attraverso il filtro dei nuovi e dei vecchi media: dai social network, al cinema, ai prodotti visivamente stereotipati delle piattaforme di streaming. Una serie di stilemi iconici marcati e ricorrenti ci richiama a questa continua omologazione: la bassissima profondità di campo, i tagli prospettici spinti, i contrasti luminosi forzati, le composizioni geometriche con ripetuti richiami alla geometria frattale, la dominanza delle curve e delle composizioni stereometriche. Da qui una seconda possibile caratterizzazione dei prodotti dell'eidomatica generativa contemporanea: la sensazione di una falsificazione piatta, in stile, rassicurante, che potremmo definire, parafrasando Umberto Eco, iporealistica: «L'Iperrealismo denuncia il fatto che la realtà, come ci siamo abituati a vederla, è effetto di una manipolazione meccanica: e quindi mette in pubblico la propria falsità programmatica. L'Iporealismo è bugiardo perché vuol farci credere che dica la verità, mentre l'Iperrealismo mette subito in chiaro il fatto che sta dicendo bugie. Questa è la grande differenza fra i due». (Eco, 1985: 78). Eco ci aiuta a sfatare il luogo comune di una virtualizzazione dell'immaginario, oggi estremizzata dalla capacità dagli strumenti computazionali, che arrivi a impedirci di distinguere il vero dal falso, facendoci prediligere la seduzione dell'artificio rispetto alla realtà. Pare invece di trovarci a una ennesima reviviscenza di un gusto per il pittoresco, subdolo e persino naïve, ancora più evidente quando si chiede alle macchine di produrre immagini urbane e paesaggistiche. Il mimetismo digitale di per sé non ci stupisce più, siamo molto più direttamente assuefatti dalla proliferazione automatizzata di un genere di suo ridondante opposto: «Il processo mimetico nei confronti della natura, poi, si inverte simmetricamente nell'idea e nella poetica del pittoresco: qui, è la realtà fisica che è giudicata bella ed esteticamente valida perché contiene o assume i caratteri di una rappresentazione; non è tanto questione di un dipinto più o meno somigliante al vero, quanto piuttosto è la natura, il vero, a essere proprio come se fosse dipinto, ad avere uno statuto pittorico» (Ugo, 1996:62).

Questo rimando speculare anamorfico fra reale e sua messa in immagine digitale da un lato veicola una memorizzazione eidetica fallace, sempre meno capace di cogliere e interpretare le origini e le operazioni che conducono all'icona finale, dall'altro innesca un cortocircuito e un'inversione fra una percezione oculocentrica dello spazio e la sua esperienza soggettiva diretta. Anche in questo preciso aspetto, come ci ricorda Macarthur, il pittoresco è dietro l'angolo: «While the picturesque is famous for the idea of forming phenomenal experience, the majority of its work lies in the simpler matter of weakening the form of the object so as to foreground subjective experience. It is a simple equation that the picturesque proposes: weaken the form of the object to foreground the forming of phenomenal experience» (Macarthur, 2007:156)

Le immagini che accompagnano questo testo, risultato di una computazione a partire da testi descrittivi volutamente laconici e provocatori, dimostrano prosaicamente la potenza evocativa e



1	2
3	4

1-4. Create con Midjourney, versione 4, prompt: 'Una nuova città futuristica, ibrida fra Milano e Roma'.



5. Creata con Midjourney, versione 4, prompt: 'Paesaggio urbano della Milano del futuro'.



6. Creata con Midjourney, versione 4, prompt: 'Una città europea del futuro'.



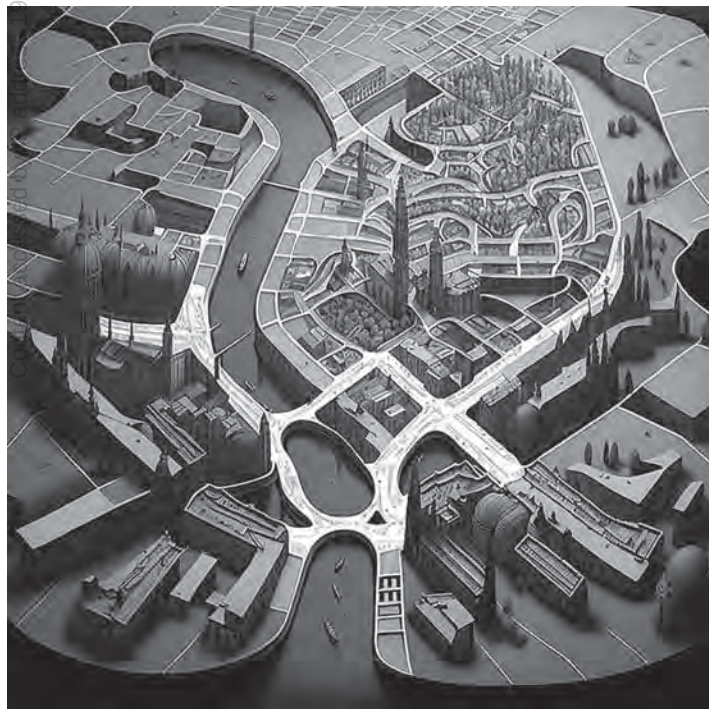
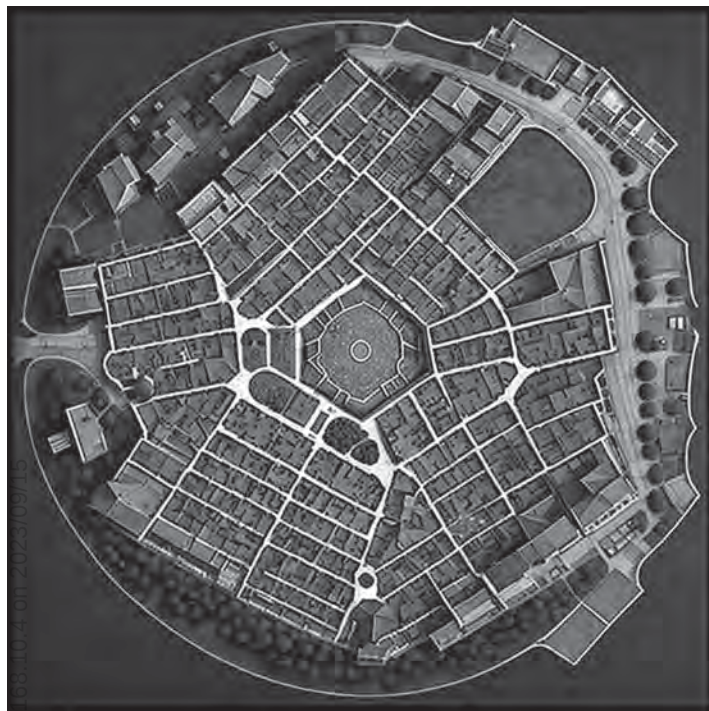
7	8
9	10

7-10. Create con Midjourney, versione 4, prompt: 'Una nuova città futuristica, ibrida fra Milano e Roma'.

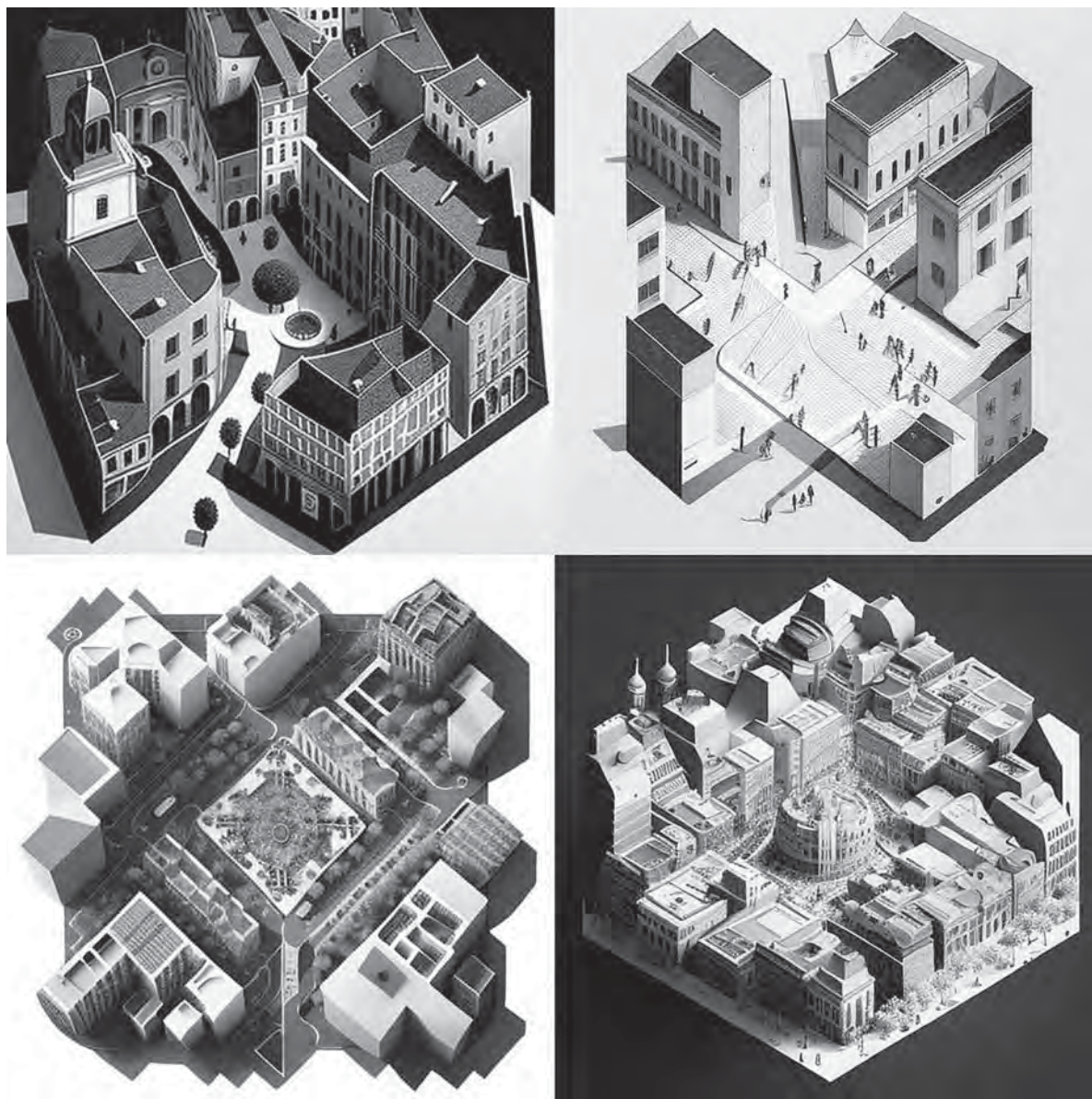


11	12	13
	14	15

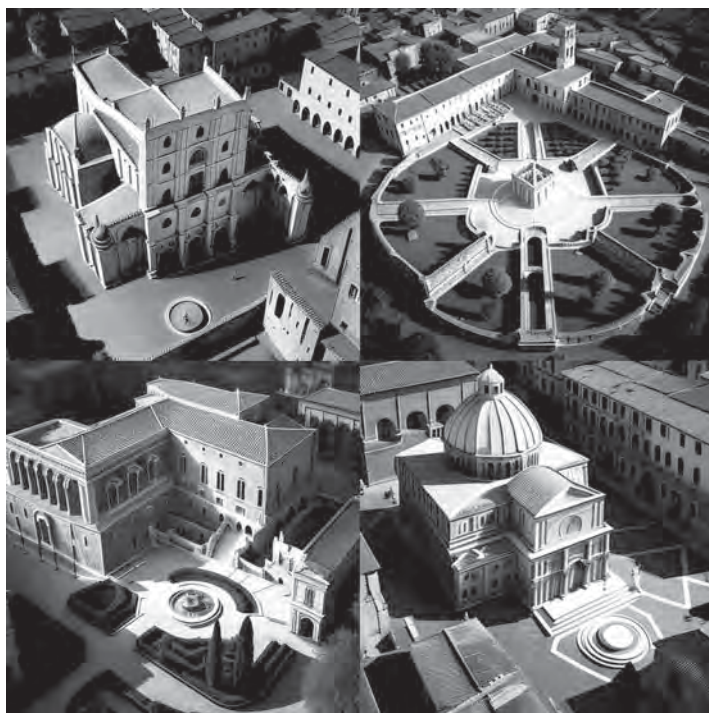
11-15. Create con Midjourney, versione 4, prompt: 'Mappa di una nuova città italiana'.







16-17. Create con Midjourney, versione 4, prompt: 'Progetto di una grande città europea'.



18	19
20	21

18-21. Create con Midjourney, versione 4, prompt: 'Paesaggio culturale italiano dall'alto'.



22. Create con Midjourney, versione 4, prompt: 'Una città italiana dall'alto'.



23. Create con Midjourney, versione 4, prompt: 'Una città italiana dall'alto'.

insieme la banalità superficiale di questa nuova generazione di simulacri. Trent'anni fa Jean Baudrillard, pur nella sua visionaria e assiomatica assertività (cfr. Latouche, 2021) costruiva i fondamenti interpretativi di una deriva digitale in cui il reale si riduce al riflesso di una manipolazione infinita, numerica, quantitativa. Un diluvio di possibili significanti, sempre più sconnessi dai loro significati: «La massa di significante fluttuante è ciò che impedisce al linguaggio di arrivare al limite delle proprie possibilità, ciò che impedisce all'uomo di esprimere tutto, e al mondo di significare tutto, di significare nella sua integralità. Ora è proprio a questo che si tende oggi attraverso le tecniche informatiche, l'Intelligenza Artificiale, ecc.: a mobilitare tutti i neuroni, tutti i sensi possibili, e simultaneamente a ridurre ogni margine, ogni vuoto interstiziale» (Baudrillard, 1992: 139).

Oltre la sua interpretazione apocalittica e volutamente parossistica (cfr. Baudrillard, 1998) di un mondo dell'immagine e degli immaginari digitali votati alla proliferazione virale, metastatica, ormai definitivamente autoreferenziale, Baudrillard ci fornisce alcuni semplici elementi interpretativi che sembrano definire con grande anticipo e precisione quello che oggi le intelligenze artificiali rendono possibile: dalla minuziosa e maniacale decostruzione della realtà nei dettagli più minuti, in una grana così fine da far perdere la relazione fra particolare e contesto (atomi, pixel senza cui non sarebbe possibile alcuna rielaborazione digitale), fino alla computazione e codificazione seriale e quantitativa, attuata attraverso azioni di simulazione, sdoppiamento, specularità e ripetizioni infinite. «I nostri sistemi complessi, metastatici, virali, votati alla sola dimensione esponenziale (non importa se quella dell'instabilità o della stabilità), all'eccentricità e alla scissiparità frattale indefinita non possono più aver fine. Votati a un intenso metabolismo, a un'intensa metastasi interna, si esauriscono in sé stessi e non hanno più destinazione, né fine, né alterità, né fatalità. Sono precisamente votati all'epidemia, alle escrescenze senza fine del frattale, e non alla reversibilità e alla risoluzione perfetta del fatale» (Baudrillard, 1992: 154). Forse la superficialità seducente e iporealistica delle immagini

eidomatiche contemporanee, prodotte da un'intelligenza artificiosa, più che artificiale, dissimula una *hybris* che in parte non sappiamo ancora capire e controllare, e che ci richiama a una ricerca accurata di nuovi, più efficaci, strumenti qualitativi di critica creativa: «Intelligence consists not only of creative conjectures but also of creative criticism. Human-style thought is based on possible explanations and error correction, a process that gradually limits what possibilities can be rationally considered. (As Sherlock Holmes said to Dr. Watson, 'When you have eliminated the impossible, whatever remains, however improbable, must be the truth:'). [...] While scientists certainly seek theories that have a high degree of empirical corroboration, as the philosopher Karl Popper noted, 'we do not seek highly probable theories but explanations; that is to say, powerful and highly improbable theories'» (Chomsky, Roberts, Watmull, 2023).

Riferimenti bibliografici

- Baudrillard J., 1992, *L'illusion de la fin*. Paris: Editions Galilée, (trad. it., 1993, *L'illusione della Fine*, Milano: Anabasi).
- Baudrillard J., 1976, *L'Echange symbolique e la mort*. Paris: Editions Gallimard (trad. it., 1979, *Lo scambio simbolico e la morte*, Milano: Feltrinelli).
- Baudrillard J., 1998, *Il Parossista indifferente*. Milano: King Kamehameha Press.
- Chomsky N., Roberts I., Watmull J., 2023, «The False Promise of Chat-GPT». *The New York Times*, online edition March 8, 2023.
- Eco U., 1985, *Sugli specchi ed altri saggi*. Milano: Bompiani.
- Indiano A., 2023, «Come l'intelligenza artificiale può creare falsi ricordi». *Wired*, edizione digitale, 17 gennaio 2023.
- Latouche S., 2021, *Quel che resta di Baudrillard*. Torino: Bollati Boringhieri.
- Macarthur J., 2007, *The Picturesque, architecture, disgust and other irregularities*. London: Routledge.
- Maldonado T., 1992, *Reale e Virtuale*. Milano: Feltrinelli.
- Ugo V., 1996, *Architettura ad Vocem*. Milano: Guerini e Associati.
- Manovich L., 2023, *Cultural Analytics. L'analisi computazionale della cultura*. Milano: Raffaello Cortina.