



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN
PHILOSOPHY, SCIENCE, COGNITION, AND SEMIOTICS (PSCS)
Ciclo 38

Settore Concorsuale: 11/C4 - ESTETICA E FILOSOFIA DEI LINGUAGGI

Settore Scientifico Disciplinare: M-FIL/05 - FILOSOFIA E TEORIA DEI LINGUAGGI

**PROCESSI DI SEMIOSI ALGORITMICA E INTELLIGENZA ARTIFICIALE.
STRUMENTI PER ORIENTARE LE POLICY IN MATERIA DI COMPETENZE
DIGITALI DA UNA PROSPETTIVA DI GENERE**

Presentata da: Flavia Politi

Coordinatore Dottorato

Luisa Lugli

Supervisore

Claudio Paolucci

Co-supervisore

Cristina Demaria

Aldo Gangemi

Esame finale anno 2026

Finanziato dall'Unione europea-

Next Generation EU,

Missione 4, Componente 2, Investimento 3.3

(D.M. 352/2022)

CUP J33C22002060009

Processi di semiosi algoritmica e Intelligenza Artificiale. Strumenti per orientare le policy in materia di competenze digitali da una prospettiva di genere.

Cosa siano queste finzioni che proponiamo in cambio delle predicazioni di strutture oggettive, è chiaro, sono i codici come istituzioni sociali e tentativi ipotetici di indicare i meccanismi degli eventi, e come i messaggi si originano dai sistemi di regole che vi presiedono. Ma opporre delle definizioni, sapute come tali, al riconoscimento virile della impossibilità di definire una origine, non è uno sfuggire all'ammissione suprema? [...]

Se l'ammissione suprema è il riconoscimento della ferita che ci consegna sin dal principio alla morte, sfuggirvi significa rimuoverla. Se è il riconoscimento del gioco, rifiutarla è ammetterla: salvo che l'ammissione anziché fermarsi su se stessa produce una poetica del gioco, e mi porta ad architettare trappole, aquiloni, fuochi d'artificio, gadget operativi manifestatamente ludici e consolatori: la scienza, i metodi, la Cultura. Ma se soltanto una di queste trappole servisse a costruire qualcosa che possa colmare il desiderio che ci muove? Non è possibile, mi risponde il pensiero ontologico, non avrai fatto altro che condannarti ancora una volta alla beffa in cui ti trovavi. Ma se soltanto una di queste trappole riuscisse a mutare di un palmo la situazione da cui sono partito? Non sarebbe successo nulla, risponde il pensiero ontologico, perché la situazione in cui ti ritrovi è ancora allo stesso numero di passi dal nulla. Non avrai spostato di un millimetro la vicinanza alla morte, che ti costituisce. Al massimo, se nel gioco avrai trovato la gioia, sarai guarito nell'accettazione dell'eterno ritorno. La coerenza di queste risposte è tale [...] che non rimarrebbe che accettarle. E tacere. Ma solo se si decidesse di continuare a muoversi nel giro delle deduzioni implicate nella domanda di partenza, ponendo la quale si era ancora fuori dal pensiero, e vi ci si entrava. E la domanda era "Chi parla"?

Umberto Eco, *La struttura assente*

Indice generale

Indice generale.....	- 3 -
Indice delle figure.....	- 5 -
Introduzione	- 7 -
0. Per una definizione di semiosi algoritmica	- 23 -
0.1 Alla base dell'emersione del senso	- 23 -
0.2 Quali competenze per le macchine dotate di linguaggio?.....	- 28 -
0.3 Allineamenti multimodali e teoria della narratività	- 31 -
0.4 Competenze negate	- 35 -
1.0 Processi di significazione o dell'IA in sé.....	- 39 -
1.1 Della competenza semiotica dei Large Language Model	- 50 -
2.0 Per una definizione dei Large Language Models.....	- 69 -
2.1 Dall'apprendimento all'inferenza: alla base dei mondi possibili.....	- 84 -
3.0 Sulla traduzione negli LLM	- 90 -
3.1 Fare ordine: interpretare, tradurre, generare.....	- 95 -
3.2 Della possibilità di operationalizzare un postulato semiotico.....	- 99 -
3.3 La natura del segno e il primato epistemologico del locale sul globale	- 102 -
3.4 Passeggiate inferenziali e ottimizzazioni	- 109 -
4.0 L'Enciclopedia echiana come virtuale intensivo	- 113 -
4.1 Dei virtuali molteplici o delle molteplicità del virtuale.	- 114 -
4.2 Differenza e differenziale	- 121 -
5.0 Dell'IA in sé tra strutture e assemblaggi	- 130 -
5.1 Lo schematismo posizionale morfodinamico e la sintassi di attrattori	- 133 -
5.2 Della semantica distribuzionale.....	- 141 -
5.3 Eterogenesi differenziale: quando una topologia dinamica incontra il virtuale enciclopedico.....	- 145 -
6.0 Una lettura post-strutturale dell'enunciazione	- 153 -
6.1 Per una topologia della soggettività.....	- 154 -
6.2 Dall'omogeneo all'eterogeneo.....	- 156 -
6.3 Identità e differenziale	- 157 -
6.4 Eterogenesi ed enunciazione.....	- 159 -
6.5 Evenemenzialità e Impersonalità.....	- 160 -

7.0 Di alcune conseguenze teoriche di capacità negate	- 163 -
8.0 Assemblaggi significanti o dell'IA per sé	- 170 -
8.1 Una premessa senza promesse. Definire l'assemblaggio.....	- 170 -
8.2 Divenire (ancora più) cyborg	- 176 -
8.3 Predicati soggettali: pensare, parlare, produrre significato	- 183 -
8.3.1 Una variazione sul tema del pensare.....	- 184 -
8.3.2 Due variazioni sul tema del parlare	- 185 -
8.3.3 Una variazione sul tema della produzione di significato	- 187 -
9.0 Filiera formativa STEM e questioni di genere ad essa connesse	- 189 -
9.1 Il genere come categoria analitica dello spazio-evento	- 189 -
9.2 La ricerca in impresa: la collaborazione con ART-ER.....	- 195 -
9.2.1 Descrizione primo intervento: Big Data and Artificial Intelligence Emilia-Romagna.....	- 198 -
9.2.2 Descrizione secondo intervento: Ragazze Digitali ER 2023	- 201 -
10. A guisa di conclusione.....	- 209 -
Bibliografia	- 212 -
Sitografia.....	- 232 -
Abstract.....	- 233 -
Italiano	- 233 -
Inglese	- 233 -

Indice delle figure

Figura 1: Modello concettuale dell'esperimento Stern-Gerlach	- 44 -
Figura 2: Immagine mentale in 3D di una rete neurale artificiale secondo François Chollet.	- 46 -
Figura 3 a) Architettura di una rete neurale, i suoi strati, la funzione di perdita e l'ottimizzatore b) Trasformazioni non lineari attivate dalla funzione ReLU	- 48 -
Figura 4: A sinistra, un feedback loop tradizionale; a destra, un feedback che risulta dalla misurazione dell'ambiente.....	- 49 -
Figura 5: Notebook "Sentiment Analysis Python".	- 55 -
Figura 6: Categorizzazione delle affermazioni prodotte da un sistema di IA generativa intesa come funzione tra "potere di selezione strategica" e "verità"	- 61 -
Figura 7: Diagramma che illustra la relazione tra le credenze del sistema di AI, le sue affermazione e il mondo.....	- 63 -
Figura 8: Sintesi visiva dell'evoluzione dei modelli Transformer	- 73 -
Figura 9: a) Neurone artificiale; b) Feed-forward neural network con un hidden layer	- 76 -
Figura 10: L'architettura modello completo del Transformer.	- 77 -
Figura 11: Attention-head view for BERT.	- 79 -
Figura 12: Model view of BERT, for same inputs as in Figure 11.....	- 80 -
Figura 13: Neuron view of BERT for layer 0, head 0 (same one depicted in Figure 11, left).-	- 80 -
Figura 14: Tipi di interpretazione.	- 92 -
Figura 15: Esempi di traduzione automatica con Babel Fish.....	- 95 -
Figura 16: Espressione e contenuto di una funzione semiotica.	- 104 -
Figura 17: Paesaggi epigenetici e loro interpretazioni.....	- 105 -
Figura 18: Punti di ottimo locale e globale	- 110 -
Figura 19: Cammino di Lévy a sinistra, moto Browniano a destra.	- 111 -
Figura 20: Varietà localmente definita e mappata su una carta singola.....	- 124 -
Figura 21: Evoluzioni di due sistemi.	- 128 -
Figura 22: Potenziali locali.	- 139 -
Figura 23: Il grafo di dipendenze sintattiche della frase “Il gatto dorme sul tappeto”.	- 140 -
Figura 24: : Relazioni sintagmatiche e paradigmatiche.	- 143 -
Figura 25: Mappatura di commensutabilità e generazione.	- 147 -
Figura 26: Il rapporto di presupposizione delle grandezze semiotiche.....	- 156 -
Figura 27: Mappa delle relazioni linguistico lessicali del termine “genere”	- 191 -
Figura 28: Organigramma e relazioni tra funzioni e comparti.....	- 196 -
Figura 29: Sintesi visuale impianto di valutazione dei risultati e di impatto.	- 198 -

Figura 30: Sintesi del quadro delle competenze digitali del personale docente ed educativo DigComp Edu.	- 200 -
Figura 31: Sintesi del quadro delle competenze digitali per la cittadinanza DigComp 2.2..	- 206 -

Introduzione

L'origine del lavoro teorico alla base di questa tesi è l'idea che esista una qualche analogia che accomuna o che permetta di leggere sinotticamente il rapporto tra prompt e modello di linguaggio di grandi dimensioni, e il rapporto tra una struttura e un virtuale intensivo. Proveremo cioè a sostenere l'idea che il prompt, ossia la richiesta in input impartita tramite interfaccia utente a un Large Language Model (LLM) ed espressa tramite linguaggio naturale, agisca come selezione o ritaglio enciclopedico locale della conoscenza appresa dal modello in fase di addestramento, conoscenza che sappiamo essere condensata nei pesi del modello. Premettiamo sin da subito che non sono i dataset in sé a rappresentare il virtuale intensivo ed enciclopedico – per una ragione che verrà spiegata nel seguito di questa introduzione e ancora in altri punti di questo elaborato (Cfr. capitolo quattro). La virtualità sta per noi nel passaggio di mediazione traduttiva in cui i valori delle unità linguistiche si ripresentano in ragione dell'intervento di un'esteriorità (il prompt umano). Da qui ci interessa ragionare sul significato della generazione dell'output testuale prodotto dal modello, ossia ci interessa chiarire quale sia lo statuto - tra ontologia ed epistemologia - del testo generato dall'agente conversazionale, nella sua relazione con la produzione di enunciati che invece appartiene all'umano e che per lungo tempo è stata ritenuta sua prerogativa assoluta. La relazione tra umano e tecnologie produttrici di testi vuole essere indagata anche attraverso le lenti delle competenze digitali implicate sia dalla prospettiva dei percorsi formativi dominio specifici nei quali si suppone che tali competenze vengano acquisite – filiera formativa STEM¹ –, sia dalla prospettiva di genere che interessa tanto lo sviluppo dei sistemi di IA in sé, quanto gli stessi percorsi formativi.

Per poter operare in questa direzione, mantenendo una consistenza e una coerenza interna al progetto di ricerca, sono necessarie alcune premesse e assunzioni che riportiamo di seguito e che inseriscono la natura ibrida del presente lavoro di ricerca nel quadro generale dei dottorati promossi dal Piano Nazionale di Ripresa e Resilienza (PNRR).

Innanzitutto, una descrizione che proceda per analogie si basa su una traduzione tra sistemi di significazione e sulla definizione di parametri di commensurabilità tra questi sistemi. In ciò il nostro lavoro è semiotico, nella misura in cui ricostruisce tentativamente catene di interpretanti che attraversano campi disciplinari. La validità scientifica di una tale operazione sarà visibile a valle del processo di scrittura; tuttavia, verranno poste delle condizioni di metodo che riteniamo possano rendere conto dell'eterogeneità informativa che sostanzia la tesi sostenuta. Vogliamo riconoscere in questa operazione il senso di quella che Paolo Fabbri definiva una “semiotica attuale”: “Il problema

¹ La scelta di tale espressione, fortemente marcata in termini economici e di mercato, verrà motivata in 9.0.

della semiotica attuale è invece quello di sostituire a queste divisioni per sostanze delle divisioni per forme organizzative, per diagrammi comuni. [...] Faccio l'esempio [...] della spazialità. È molto probabile che i modelli fondamentali del linguaggio siano di tipo spaziale [...]. Non è possibile confrontare parole e sillabe, da un lato, con architravi e colonne, dall'altro; è possibile invece paragonare le forme spaziali usate dalla semantica linguistica con quelle prese in carico dalla semantica architettonica. In questo modo, architettura e linguaggio diventano traducibili". (Fabbri 1998, in Paolucci 2010, p. 256).

Procedere per analogia, per noi, è una scelta di metodo con funzione accrescitiva della conoscenza. Accade che si ha a che fare con una serie di fenomeni, fatti, nozioni, comportamenti che sono separati per via gnoseologica o per via ontologica; tra questi fenomeni, fatti, nozioni, comportamenti, sussistono delle relazioni interne tra le loro parti che sono in qualche modo corrispondenti alle relazioni interne tra le parti di un altro fenomeno. Detto altrimenti ciascuno di questi fenomeni ha una commensurabilità con le proprietà di un altro fenomeno che nella trattazione e analisi tradizionale non verrebbe altrimenti presa in considerazione. Si tratta di *un'analogia di proporzionalità* esprimibile nella formula $A: X_A = B: X_B$. Questa struttura matematica è stato il punto di partenza della nostra ricerca, quando abbiamo impostato il rapporto tra Enciclopedia e s-codice con il rapporto tra LLM e un algoritmo genericamente inteso. Vedremo meglio questa analogia in 1.1, ma accompagnerà diversi passaggi di questo elaborato, ogni qual volta sarà necessaria una trasformazione tra sistemi di segni e non solo una loro comunicazione (un esempio sono le metaeuristiche di addestramento delle reti neurali artificiali che vedremo in 3.4: qui la pertinenza dell'analogia di proporzionalità riposa sull'idea che le passeggiate inferenziali nell'enciclopedia individuale presentino delle proprietà simili ai processi di ottimizzazione che gli algoritmi di training eseguono nei *search space* dei LLM).

Così Eco descriveva il portato e il valore conoscitivo dell'analogia, già riconosciuto da Aristotele: "Si riconsideri l'esempio dei pirati detti fornitori. Anzitutto s'instaura una analogia a quattro termini: i pirati stanno al trasporto delle merci rubate come i fornitori stanno al trasporto delle merci acquistate. Dopo, l'impressione che si individui un genere x di cui sono specie sia i pirati che i fornitori, è conseguente all'operazione analogica. Essa di fatto prende due sememi indipendenti e vi individua una proprietà comune (di essere trasportatori di merci). Solo a metafora compresa si potrebbe dire che pirati e fornitori appartengono (inopinatamente) allo stesso genere, ovvero allo stesso insieme. La proprietà comune, sorprendentemente messa in rilievo, diventa genere comune." (Eco 2004, p. 105). E ancora: "Ma nell'analogia aristotelica si scopre l'identità di proprietà tra due cose le cui

proprietà sono note (la scoperta riguarda la relazione inedita che si pone tra due cose note)” (Ivi, p. 107). Il lettore e la lettrice valuteranno l’adeguatezza Foucault di questa operazione.

Preme inoltre sottolineare che il percorso di ricerca dottorale è stato caratterizzato, di per sé, da eventi e contenuti non circoscrivibili esclusivamente a singole tradizioni scientifiche e/o scuole di affiliazione: il Dipartimento di Filosofia (FILO) dell’Università di Bologna, l’area Transizione Digitale di ART-ER S.Cons.P.A. e il Centre d'analyse et de mathématique sociales (CAMS) de l’École des hautes études en sciences sociales (EHESS) sono stati i luoghi in cui la ricerca ha radicato i suoi strumenti di analisi, ha definito le sue domande guida e ha individuato le traiettorie di avanzamento.

Questi luoghi si configurano come reti organizzative e umane che hanno consentito di valorizzare la pluralità di sguardi e posizionamenti sulla materia trattata, insieme agli interessi teorici specifici di chi scrive. I programmi innovativi di dottorato promossi dal Piano Nazionale di Ripresa e Resilienza, con il contributo e il coinvolgimento delle imprese, rispondono al bisogno di “migliorare le competenze in settori centrali per lo sviluppo”²; ciò implica che la stesura della tesi e il lavoro di ricerca siano in grado di dialogare con priorità e istanze anche esterne al contesto accademico, in una logica di scambio di conoscenze e competenze utili alla crescita di quello che viene denominato “l’ecosistema regionale dell’innovazione”.

Da questo punto di vista, particolarmente sfidante risulta essere la selezione di un registro linguistico e uno stile testuale adeguato alla pluralità dei profili di competenze di potenziali lettori empirici e lettrici empiriche. Si tratta quindi di distillare quello che Umberto Eco chiamava “Lettore Modello il cui profilo intellettuale è determinato solo dal tipo di operazioni interpretative che si suppone (e si esige) che egli sappia compiere: riconoscere similarità, prendere in considerazione certi giochi...” (Eco 1979, p. 87)³, non un individuo in carne e ossa, ma una strategia testuale che, assieme all’Autore Modello renda possibile il funzionamento e l’interpretazione del testo. Accanto alle operazioni interpretative, infatti, vanno prese poi in considerazione le operazioni cooperative e “i doveri filologici”, cioè “recuperare con la massima attenzione possibile i codici dell’emittente”. A un livello di scrittura, questo è stato l’approccio adottato in particolare in fase di analisi e studio di

² <https://www.italiadomani.gov.it/content/sogei-ng/it/it/Interventi/investimenti/Assegnazione-di-borse-di-dottorato-in-collaborazione-con-le-imprese-e-promozione-dell'assunzione-dei-ricercatori-dalle-imprese.html> (ultimo accesso: 15 febbraio 2026)

³ Eco, U. (1979). *Lector in fabula*. La Nave di Teseo (ed. 2020)

contenuti testuali provenienti da ambiti disciplinari discontinui rispetto al profilo di competenze di chi scrive.

Sarà quindi all'intersezione tra macroaree di ricerca che questa tesi proverà a collocarsi attingendo alla semiotica interpretativa e cognitiva, alla filosofia del linguaggio e all'epistemologia, alle matematiche e alla computer science, quest'ultima luogo naturale del campo di studi che va sotto il termine ombrello Intelligenza Artificiale (IA).

Strettamente connessa alla prospettiva transdisciplinare va rilevata un'ulteriore difficoltà di natura comunicativa sottesa ad affrontare un discorso sui sistemi di IA generativa che fattualmente hanno rotto il tetto di cristallo della riproduzione del linguaggio naturale, rispetto ad altri sistemi di Intelligenza Artificiale integrati ormai in molteplici oggetti tecnici quotidianamente utilizzati. Fino all'avvento dell'architettura Transformer - vedremo come all'inizio anche questa non ricevette le necessarie attenzioni – e più propriamente alla commercializzazione dell'agente conversazionale che l'utenza più diversificata conosce sotto il nome di ChatGPT⁴, l'attenzione a livello di cultura di massa era ridotta, il dibattito c'era ma per lo più circoscritto all'ambito scientifico e dunque veniva trattato solo da esperti ed esperte di settore. La diffusione al pubblico dei grandi modelli di linguaggio e delle loro interfacce conversazionali (Fig. 1) fa entrare nel dibattito pubblico generalista la promessa di macchine intelligenti dotate di linguaggio, con una serie di conseguenze che rendono complessa una presa di parola sottratta a spinte polarizzanti dove forte è la contaminazione tra opinione (interessata e indirizzata) e discorso scientifico.⁵

Fornitore	LLM	Interfaccia chatbot ufficiale
OpenAI	GPT family	ChatGPT (web, iOS, Android)

⁴ Uno sguardo demografico aggiornato sulle statistiche di utilizzo di ChatGPT è disponibile in Duarte, F. (2025), *Number of ChatGPT Users* (July 2025) a questo link: <https://explodingtopics.com/blog/chatgpt-users#how-many> (ultimo accesso: 15 febbraio 2026).

⁵ Per analisi quantitative basate su metriche di polarizzazione del dibattito all'interno dei social media e delle piattaforme online sui Large Language Models si vedano a titolo esemplificativo: Qi, W., Pan, J., Lyu, H., & Luo, J. (2024). Excitements and concerns in the post-ChatGPT era: Deciphering public perception of AI through social media analysis. *Telematics and Informatics*, 92, 102158; Alipour, S., Galeazzi, A., Sangiorgio, E., Avalle, M., Bojic, L., Cinelli, M., & Quattrocioni, W. (2024). Cross-platform social dynamics: An analysis of ChatGPT and COVID-19 vaccine conversations. *Scientific Reports*, 14(1), 2789. Per un'analisi empirica di taglio sociologico applicata al contesto della stampa italiana prima e dopo ChatGPT si veda: Degli Esposti, P., Bonazzi, M., & Spillare, S. (2024). AI Imaginaries and Narratives in the Italian Public Discourse: The Impact of ChatGPT. *IM@GO*, 23, 132–147. Non entriamo nel merito della costruzione di questi strumenti di analisi poiché eccede gli intenti di questo elaborato; il riferimento a questi studi, tuttavia, definisce il quadro al di fuori del quale vorremmo provare a collocare il nostro intervento.

Fornitore	LLM	Interfaccia chatbot ufficiale
Anthropic	Claude	Claude.ai + app
Google	Gemini	Gemini (web + app)
Microsoft	Copilot	Copilot (web, Windows/Edge, mobile)
Meta	Llama	Meta AI (WhatsApp, Instagram, Messenger, Facebook)
Mistral	Mistral	Le Chat (web + app)
Perplexity	Perplexity	Perplexity (web + app)
xAI	Grok	Grok (web, app, X/Twitter)
DeepSeek	R1, V3, ecc....	DeepSeek Chat (web; app ⁶)
Baidu	ERNIE	ERNIE Bot

Figure 1: Tabella generata con ChatGPT modello 5 auto. Prompt inviato il 19 agosto 2025: “Crea una tabella contenente i LLM attualmente in commercio e le interfacce chatbot di ciascuno, specificando se sono disponibili in Italia gratuitamente o a pagamento”. La tabella originaria prodotta dal chatbot, modificata per le presenti esigenze, riuniva in un’unica colonna Fornitore e modello di LLM; presentava inoltre altre 3 colonne riferite all’accessibilità del modello in Italia, alla disponibilità di piani gratuiti e le tipologie di piani a pagamento. Sono state necessarie delle verifiche e delle correzioni sui contenuti prodotti dal modello di linguaggio.

Fatta luce su questi aspetti di contesto, iniziamo ora a delineare a quale livello teorico vuole intervenire il presente elaborato.

Come ha affermato Beatrice Fazi⁷: “Although researchers have been experimenting with generative computing programmes since the 1960s, in the present technocultural context the very idea of ‘machines that create’ is taken to signal an inflection point, as heralding a new, contemporary technological paradigm and representing a disruptive phase in the history of computing, altering the way in which society and economy operate” (Fazi 2024, 2-3).

Un pregio della posizione di Fazi è quello di non ignorare il portato trasformativo che accompagna lo sviluppo delle nuove tecniche di machine learning, senza tuttavia assumere totalmente e acriticamente lo sfondo socio-tecnico in cui si inseriscono. In nota all’articolo già menzionato commenta infatti:

⁶ Si veda il provvedimento emesso dal Garante italiano per la protezione dei dati personali il 30 gennaio 2025, disponibile al seguente indirizzo: https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10098477?utm_source=chatgpt.com (ultimo accesso: 15 febbraio 2026)

⁷ Fazi, M. B. (2024). Machines That Create: Contingent Computation and Generative AI. *Media Theory*, 8(2). <https://doi.org/10.70064/mt.v8i2.1115>

“I am interested in the speculative avenues that these systems open up for philosophy although also aware of the commercial propaganda that surrounds them. I believe it is possible to be curious about the philosophical implications of these systems while also remaining critical of some of the narratives and uses that accompany them, specifically in relation to how technological apparatuses comply with and also further the corporate aims and profit-oriented goals of post-industrial, informational society” (Ivi, p. 12).

Se l’accento di Fazi è posto sulle sollecitazioni che i sistemi di IA generativa stimolano nel pensiero filosofico verso una definizione di cosa sia il pensiero, la creazione o la creatività a livello computazionale, il nostro sguardo vuole invece cogliere la specificità semiotica di questi strumenti e delle formazioni discorsive e concettuali che cercano di renderne conto.

Negli anni che hanno fatto seguito alla commercializzazione massiva degli attuali sistemi generativi e sintografici, il campo semiotico ha individuato specifiche linee di analisi che rappresentano i contorni di una semiotica dell’Intelligenza Artificiale. Già agli inizi del 2023 Massimo Leone⁸ indicava tre strade percorribili e corrispondenti ai compiti di una semiotica dell’IA:

1. Lo studio degli artefatti simulativi, cioè degli sforzi di simulare il comportamento intelligente umano, il piano dell’espressione dell’intelligenza naturale. Ma non solo, poiché come egli afferma: “This simulation can take place at the level of expression, at that of content, or at both”. A ciascun piano competono i propri formanti di riferimento, ad esempio “the inorganic reproduction of signs that humans associate with intelligence” per il piano dell’espressione, dove il target di analisi diviene lo studio della retorica dell’intelligenza artificiale, cioè “the expressive signs that simulate an intelligent behaviour or device”. Sul piano del contenuto, l’oggetto di interesse sarà invece “the inorganic reproduction of meaning that humans associate with intelligence”, un significato che Leone considera comportamentale, legato cioè all’effetto di intelligenza veicolato agli occhi umani per il tramite di un comportamento intelligente da parte della macchina.

2. Lo studio delle ideologie di simulazione, cioè la messa in prospettiva storica dei sistemi semiotici umani sui quali si basano le tecnologie del significato, alla ricerca delle ideologie implicite che accompagnano lo sviluppo e la diffusione di nuovi oggetti tecnici nel mondo digitale. A questo livello, le ideologie che danno valore all’Intelligenza Artificiale coinvolgono la definizione di cosa

⁸ Leone, M. (2023a). The main tasks of a semiotics of artificial intelligence. *Language and Semiotic Studies*, 9(1), 1-13. <https://doi.org/10.1515/lass-2022-0006>

sia l'intelligenza, come dovrebbe funzionare e quale rapporto intrattiene col mondo in termini di risultati prodotti. Per Leone "Each culture and each historical epoch are characterized by the specific semiotic modalities that they adopt in the production of the fake; artificial intelligence is becoming the chief modality in the present-day digital production of fake".

3. Infine l'attenzione a specifici casi studio tra i quali Leone porta come esempio illustrativo quello della riproduzione dei volti da parte dei sistemi di IA, in particolare le reti generative avversarie basate su un algoritmo sviluppato da Goodfellow nel 2014 e descritto nell'articolo dal titolo "Generative Adversarial Nets". Come funzionano queste reti da un punto di vista semiotico? Dice Leone "Two main actants appear in the abstract architecture of GANs. The first is an actant generator that examines a data configuration and produces a text which could be issued from this very configuration; the second is an actant discriminator that examines the text thus produced and evaluates whether it comes from the data configuration or from the generating actant. From an epistemic point of view, therefore, the generating actant aims 'to make appear' and therefore 'to pass as true' what is not, while the discriminatory actant aims 'to make appear' and therefore 'to unmask as false' what is not true".

Più recentemente Dondero et al. (2025)⁹ hanno dedicato uno *special issue* della rivista *Semiotica* a tre aspetti chiave di una Semiotica dell'IA, più indirizzata verso i modelli generativi: il piano dell'enunciazione, quello dell'agency e quello della creatività. Nell'editoriale di apertura si riconosce come la Semiotica abbia lavorato sull'Intelligenza Artificiale sin dagli anni Sessanta, ma vengono individuati alcuni momenti di svolta rappresentati prima dall'interesse rivolto verso l'Image Processing nel 2016 e successivamente verso l'Intelligenza Artificiale generativa a partire dal 2022. Ancora, rilevano come uno spazio di riflessione semiotica a livello europeo sul tema dell'IA generativa sia rappresentato dal *Séminaire international de sémiotique à Paris*, coordinato da Maria Giulia Dondero e Juan Alonso Aldama, nel quale per due annualità consecutive, 2023/24 e 2024/25, sono stati offerti contributi multidisciplinari su diversi temi centrali per un'analisi delle tecnologie che producono *testi*. In particolare, nella prima annualità il focus è stato indirizzato al problema dell'enunciazione e alle passioni che attraversano il dominio semiotico aperto dall'IA. Nell'ultima annualità invece lo spazio di confronto è stato costruito a partire dalle sfide poste in essere dalla traduzione all'interno dei modelli computazionali di linguaggio, traduzione intesa come passaggi

⁹ Dondero, M., Alonso Aldama, J. & Leone, M. (2025). Aspects of AI semiotics: enunciation, agency, and creativity. *Semiotica*, 2025(262), 1-3. <https://doi.org/10.1515/sem-2025-0014>

“entre prompts humains et prompts révisés par la machine, entre prompts en langage naturel et descriptions textuelles d’images”¹⁰.

Volendo tracciare alcune delle linee interpretative che emergono dal numero speciale della rivista di settore *Semiotica* e che quindi restituiscono il piano di articolazione più attuale che il campo semiotico sta portando avanti rispetto ai sistemi di IA generativa, possiamo individuare almeno due campi tensivi utili all’inquadramento del presente lavoro di ricerca.

Innanzitutto, il problema dell’enunciazione apre a diversi posizionamenti, a livello teorico, che attingono a una legittimazione nella quantità e nella qualità di agency si ritiene possa essere attribuita ai sistemi macchinici, così come alla loro proprietà creatrice. Si può concepire che la macchina intelligente enunci nella misura in cui non si ritenga che l’agency dell’atto enunciativo, cioè del passaggio di conversione tra una virtualità – la *langue* saussuriana per Dondero (2025)¹¹ e, con una sfumatura differente, l’enciclopedia echiana per Paolucci (2025a)¹² – e la sua attualizzazione nell’atto di *parole*, appartenga in via esclusiva al soggetto che produce quell’enunciato speciale: l’io (enunciatore) che dice io (soggetto linguistico dell’enunciato).

Il virtuale in Saussure è la *langue*, che è il fatto sociale del linguaggio, ciò che residua se sottraiamo al linguaggio l’individuale e l’accidentale, ossia la *parole*. Dice Saussure: “Se potessimo abbracciare la somma delle immagini verbali immagazzinate in tutti gli individui, toccheremmo il legame sociale che costituisce la lingua (*langue*). Questa è un tesoro depositato dalla pratica della *parole* nei soggetti appartenenti a una stessa comunità, un sistema grammaticale esistente virtualmente in ciascun cervello o, più esattamente, nel cervello di un insieme di individui, dato che la lingua non è completa in nessun singolo individuo, ma esiste perfettamente soltanto nella massa” (Saussure 1922, p.23).¹³

Al contrario il virtuale enciclopedico di Eco viene presentato dallo stesso autore come “un’idea regolativa” che “serve a individuare porzioni di enciclopedia attivabili nella misura in cui queste

¹⁰ <https://www.fmsch.fr/projets/seminaire-international-de-semiotique-paris-3>

¹¹ Dondero, M. (2025). Semiotics of artificial intelligence: enunciative praxis in image analysis and generation. *Semiotica*, 2025(262), 111-146. <https://doi.org/10.1515/sem-2024-0195>

¹² Paolucci, C. (2025). The myth of meaning: generative AI as language-endowed machines and the machinic essence of the human being. *Semiotica*, 2025(262), 5-23. <https://doi.org/10.1515/sem-2024-0204>

¹³ Saussure, F. d. (1922), *Cours de linguistique Générale*. Paris: Éditions Payot; tr. it. De Mauro, T. (a cura di) Corso di linguistica generale. Roma-Bari: Editori Laterza (2003).

servono a costruire gerarchie provvisorie, o reti maneggiabili, ai fini di interpretare e di spiegare la interpretabilità di certe porzioni di discorso” (Eco 2007, p. 68).

Dettaglieremo in modo sistematico questo concetto in 0.2, e ancora più approfonditamente in 4.0. Paolucci definisce l’enciclopedia come un *effetto a priori*: “Effetto, in quanto prodotto di tutti gli atti di enunciazione che essa registra; a priori – in senso materiale e storico – in quanto essa rappresenta localmente la condizione di possibilità di nuovi atti di enunciazione. L’enciclopedia è al contempo strutturata dall’enunciazione passata e strutturante l’enunciazione futura” (Paolucci 2020, p. 95). Questo doppio statuto di effettività temporale dell’enciclopedia e di contemporanea attestazione, registrazione e condizione di possibilità è per noi fondamentale al fine di poter impiegare questo costruito concettuale all’interno di una teoria della semiosi algoritmica. Approfondiremo questo aspetto nelle pagine seguenti, perché costituisce parte dell’ossatura della nostra argomentazione e di cui siamo debitrice alla teoria dell’enunciazione impersonale ed evenemenziale consegnataci da Paolucci.¹⁴

Se concepiamo l’agency come distribuita tra le grandezze semiotiche convocate all’interno dell’enunciato e che non necessariamente presentano uno statuto soggettale o immanente al piano del linguaggio, ma che rimandano alle stratificazioni del linguaggio teorizzate da Hjelmslev – per quanto più pertinenti alla *parole* –, allora le posizioni di soggetto aperte dall’enunciazione all’interno dell’enunciato possono essere anche occupate da macchine variamente intelligenti che producono questi stessi enunciati. Certo, va innanzitutto chiarito il concetto di *agency* nelle varie tassonomie in cui è articolato¹⁵, ma da un punto di vista semiotico la posizione di Valle (2025)¹⁶ ci sembra saggia

¹⁴ Paolucci, C. (2020), *Persona. Soggettività nel linguaggio e semiotica dell’enunciazione*. Milano: Bompiani.

¹⁵ Si veda Floridi, L. (2025b). AI as Agency without Intelligence: On Artificial Intelligence as a New Form of Artificial Agency and the Multiple Realisability of Agency Thesis. *Philosophy & Technology*, 38(1), 1–27. Nell’articolo Floridi identifica tre osservabili epistemiche concorrenti alla formazione di un modello di agency: l’interattività, l’autonomia e l’adattabilità. Rimandiamo al testo per approfondimenti, ma questi tre criteri di identificazioni sono considerati sufficientemente astratti per adeguarsi a diverse forme di agency che per Floridi coincidono con: l’agency naturale, biologica, animale sociale, artefattuale umana individuale, umana sociale e artificiale (o dell’IA). Ciascuna di queste forme di agency possiede un certo grado di attivazione delle tre osservabili, di cui vengono forniti esempi concreti ai quali si rimanda.

¹⁶ Valle, A. (2025). From grammar to text: A semiotic perspective on a paradigm shifts in computation and its usages. *Semiotica*. <https://doi.org/10.1515/sem-2024-0187>. Nello specifico, Valle ancora, per il tramite della teoria cibernetica, l’agency alla teleologia qui intesa come controllo per feedback: “Teleology is synonymous with purpose controlled by feedback” (Rosenblueth et al. 1943: 23), in particular with negative feedback as an adjusting, tuning operation, like the one done by the optimization function in a NN. Here Rosenblueth, Wiener and Bigelow propose a clear definition of subjectivity that seems interesting for semiotics. It is based on the idea of purpose-directed agency. It is an operational, formal, and thus very general definition” (pp. 55–56).

tenere insieme il funzionamento della macchina (nello specifico le reti neurali artificiali) e l'agire per uno scopo, l'agire orientato al valore, oltre l'intenzionalità di stampo fenomenologico. Dice, infatti, Valle: "Indeed, agency can be approached from multiple perspectives, e.g., philosophical, phenomenological, anthropological. Hence it is a particularly interesting terrain of dialogue (which has not yet occurred) between AI and semiotics. As a cursory example, a computational agent can be characterized in terms of, e.g., plans, beliefs, knowledge (Poole and Mackworth 2023). On the other hand, in semiotics, the notion of "actant" (Greimas and Courtés 1979) concerns precisely an idea of subjectivity oriented towards value; the issue of plans is discussed in terms of narrative programs; the aspect of beliefs takes various declinations, e.g., in terms of value or passion; knowledge can be thought of as competence" (Ivi, pp. 2-3). Questo modo di intendere l'agency apre inoltre a quella che più avanti definiremo una strategia discorsiva, ossia quella di un farsi soggettività dei sistemi di AI.

Proseguendo, l'intenzionalità è in D'Armenio (2025) precisamente l'assenza del tratto marcato dell'intelligenza macchinica. Dopo aver richiamato la definizione di semiotica proposta da Basso Fossali (2017), ossia la scienza di tutte le mediazioni della gestione sociale del significato articolata nelle quattro sfere del senso – fenomenico-percettivo, linguistico, istituzionale e tecnologico – D'Armenio ne rilegge il funzionamento alla luce dell'impatto che i sistemi di IA hanno "sulle dinamiche semiotiche relative a ciascuna di queste sfere" (trad. nostra). Ciò che più rileva qui è il legame costitutivo tra IA e prassi enunciativa: "car leur travail est précisément un travail de convocation, d'actualisation et de réalisation des possibilités potentielles et virtuelles résultant de la sédimentation des énoncés humains organisés en bases de données structurées. Cependant, le fonctionnement de l'IA pose également des questions théoriques à la sémiotique, notamment en ce qui concerne l'intentionnalité qui habite implicitement la théorie de l'énonciation. Si les IA produisent des énoncés, elles ne possèdent pas, à l'état actuel, aucune forme d'intentionnalité". In questo senso egli parla, limitatamente al campo dell'IA generativa visuale, di "entités co-énonçantes, dépourvues d'intentionnalité et d'initiative, qui produisent pourtant des énoncés en collaboration avec un opérateur humain et sur la base d'archives hautement structurées et reconfigurables de données". L'autonomia enunciativa è erosa dall'assenza di intenzionalità dell'atto stesso di enunciazione. Ci domandiamo, di conseguenza, se sia possibile conservare la tenuta concettuale di una teoria dell'enunciazione di stampo benvenistiano, investendo del ruolo di co-parlante un'entità macchinica. Delle due l'una: o l'enunciazione non ha nulla a che fare con un principio sintetico *a posteriori* – e non *a priori* come Kant voleva – dell'esperienza in grado di farsi soggetto attraverso il linguaggio e

che ricava il “fondamento della soggettività” a partire dallo “status linguistico della persona”¹⁷ e dunque il fatto che un sistema di IA generativa dica “io” non ne fa una macchina enunciante in prima persona, ma solo un’entità co-enunciante; oppure l’enunciazione e l’agency hanno a che fare solo in via secondaria con il linguaggio e primariamente si fondano su assemblaggi o collettivi di istanze enuncianti, la cui natura è eterogenea e differenziale.

Infine, come in un gradiente scalare discendente, a livello minimo di decrescita non tanto dell’agency della macchina di cui viene riconosciuta l’autonomia relativa e il potenziale creativo, ma a livello di presa in carico enunciativa, Colas-Blaise (2025)¹⁸ propone una “séquence énonciative machinique” in cui il soggetto di enunciazione resta l’umano a monte e a valle del processo enunciativo – e non potrebbe essere altrimenti se si assume la prospettiva simulacrale dell’enunciazione¹⁹ – e le istanze macchiniche vanno a occupare la fase mediana della “textualisation numérique”, possibile in virtù di un’operazione di débrayage tra l’istanza enunciante umana, la quale determina le condizioni di possibilità stesse dell’enunciazione a monte e la convalida metalinguistica del contenuto a valle, e la macchina intelligente sulla base di un “approche pluri-instancielle distributive”. Per Colas-Blaise la relazione umano-macchina è una deiscenza, la lacerazione provocata da una ferita che, tuttavia, a nostro giudizio, non può che essere consustanziale all’enunciazione umana che è in realtà macchinica come pleonasma. In 7.0 proveremo a dare ragione di questa posizione.

Una seconda linea interpretativa, di base strettamente connessa alla teoria dell’enunciazione e sopra già menzionata, riguarda il ruolo centrale giocato dagli archivi e dai dataset di addestramento in una semiotica dell’IA, come evidenziato da Dondero (2025) e D’Armenio (2025). Dondero analizza infatti il rapporto tra immagini e database sia nei termini dell’analisi computazionale delle immagini, sia nei termini della loro generazione. Il database è lo strumento primo, accanto ai modelli algoritmici, per comprendere il funzionamento della prassi enunciativa dei linguaggi visuali odierni, in quanto permette di trattare la virtualizzazione non più come “an abstract concept” senza possibilità

¹⁷ Riprendiamo da Paolucci la definizione di soggetto in Benveniste: “Con “soggetto”, infatti, Benveniste intende «l’unità psichica che trascende la totalità delle esperienze vissute che essa riunisce, e che assicura il permanere della coscienza. Noi riteniamo che questa “soggettività”, che la si consideri da un punto di vista fenomenologico o psicologico, non importa, non è altro che l’emergere nell’essere di una proprietà fondamentale del linguaggio. È “ego” che dice “ego”. In ciò troviamo il fondamento della “soggettività”, che si determina attraverso lo status linguistico della persona»” (Paolucci 2020, p. 32).

¹⁸ Colas-Blaise, M. (2025). La machine crée, mais énonce-t-elle? Le computationnel et le digital mis en débat. *Semiotica*, 2025(262), 147-187. <https://doi.org/10.1515/sem-2024-0188>.

¹⁹ Per un confronto sull’enunciazione come atto di mediazione e sull’enunciazione come atto di scissione si veda Paolucci (2020).

di testing od operazionalizzazione. Le banche dati possono infatti essere considerate “as the concretization and totalization *in a same place* of all major visual cultural productions (and of their formal characteristics)”, ovvero concretizzano il passaggio tra sedimentazione e innovazione alla base dello studio delle trasformazioni culturali, su cui si fonda la stessa prassi enunciativa (Fontanille 1998; 2003).²⁰ Dondero mette in luce come il database sia una riserva di forme già enunciate che sono rese tra loro commensurabili e manipolabili in quanto “translated into lists of numbers”, ossia rappresentazioni vettoriali “that define the coordinates of the digital objects encoded and embedded in latent space and their relations of distance and proximity within it”, citando Somaini (2023).²¹

D’Armenio (2025) rinforza quest’idea del dataset come un tipo particolare di archivio e come ciò che consente all’IA di convocare, attualizzare, realizzare in nuovi enunciati visivi gli enunciati umani culturalmente sedimentati e organizzati in database. Ancora, l’archivio viene presentato secondo quattro accezioni, di cui tre più classiche e la quarta dall’autore aggiunta proprio in relazione alla trasformazione dell’oggetto all’interno dei sistemi di IA generativa. In quest’ultimo senso, “une base de données n’est pas simplement une collection d’archives, mais une archive d’archives”(Ivi p. 221), mantenendo le prime tre accezioni di archivio – patrimonio da trasmettere, risorsa digitale da manipolare ed effetto di senso– a cui si aggiunge l’idea di “archives d’opérations” contenenti non solo “d’images de chats associées aux descriptions correspondantes, mais d’images complexes de chats associées à des descriptions articulées, à des récits et à des jugements esthétiques”(Ivi, p.222).

Infine, seppur in forma incidentale, ma per noi molto pertinente, l’intervento di Basso Fossali (2025)²² chiama in causa “des archives réorganisées autour des intelligences artificielles” (Ivi, p. 79) come ipotesi di una semiotica di terzo ordine, nella quale si verifichi uno sdoppiamento – con sviluppi duali autonomi sul modello dei mondi possibili – delle semiosfere culturali. Più precisamente, partendo dal presupposto che la semiotica, come dice lo stesso Basso Fossali, non è scienza delle rappresentazioni o scienza delle presupposizioni logiche, ma scienza delle mediazioni, si può ritenere che ogni cultura richieda almeno due ordini semiotici, dunque due livelli di articolazioni delle mediazioni attraverso circuiti della significazione. Una semiotica di primo ordine prenderà in considerazione le rimediazioni tra individuazione sensibile dei contenuti delle manifestazioni espressive, nel passaggio a una percezione di per sé già semioticamente orientata in termini di

²⁰ Fontanille, J. (2003), *Semiotics of discourse*, Berlin, Peter Lang; Fontanille, J. & Zilberberg, C. (1998), *Tension et signification*, Liège, Mardaga.

²¹ Somaini, A. (2023), *Algorithmic images: Artificial Intelligence and visual culture*, Grey Room 93. 74–115. https://doi.org/10.1162/grey_a_00383.

²² Basso Fossali, P. (2025). Rationalités correctives et intelligence artificielle assistée : les doubles contraintes des humanités numériques. *Semiotica*, 2025(262), 71-109. <https://doi.org/10.1515/sem-2024-0189>

désenchantement o di *simulation*. Una semiotica di secondo ordine dove la rimediazione è di natura epilinguistica nella forma di un ripiegamento, dunque di una re-immissione dei valori istituiti per autoriflessività; a questo livello è evidentemente implicata l'intera cultura, il che permette di definire l'identità culturale nella sua dinamicità irriducibile a un'identità irrelata e nella dialettica tra piano di immanenza e trascendenza delle forme di vita culturali. Ipotizzare un terzo ordine semiotico, per Basso Fossali significa “imaginer d'accueillir les Intelligences Artificielles comme des épicentres de reproduction énonciative dans une écologie sémiotique dans laquelle on pourra compter alors sur des médiations ultérieures: une sémiosphère augmentée” (Ivi, p.78). Le mediazioni tra i circuiti di significazione possono contare su “nuovi terreni di significazione” (Ibidem, trad. nostra) che sono direttamente esperiti dall'utente; mentre rispetto all'ambiente nel quale può darsi l'emersione stessa dell'Intelligenza Artificiale, Basso Fossali parla di “sémiosphère artificielle”.

A questo livello di analisi, considerare il passaggio tra virtuale-attuale alla base della prassi enunciativa come *operazionalizzato* nel rapporto tra dataset/archivio e generazione di nuove occorrenze testuali e visuali ci sembra necessiti di un approfondimento in termini di quale sia la caratterizzazione della virtualità semiotica e cosa possa *stare per* il virtuale nei modelli artificiali di linguaggio. Stiamo attuando come già detto un processo di traduzione tramite catene di interpretanti, a partire dal fatto degno di nota che i sistemi di IA producono a tutti gli effetti enunciati. A nostro modo di intendere, considerare sovrapponibile il concetto di un virtuale-archivio, che è omogeneo nella natura delle attestazioni che lo compongono, con il funzionamento di un Large Language Model nelle sue diverse fasi operative - dove più composita appare la natura dei dati e delle pratiche che lo alimentano: dati di pre-training, dati di fine-tuning, tecniche di continual learning, instruction tuning for learning e così via – può portare a una riduzione della complessità. Se il database sta per la *langue* saussuriana, è lo stesso Saussure a dirci che: “Mentre il linguaggio è eterogeneo, la lingua così delimitata [“oggetto ben definito nell'insieme eteroclitico dei fatti di linguaggio” e “oggetto di natura concreta”] è di *natura omogenea*: è un sistema di segni in cui essenziale è soltanto l'unione del senso e dell'immagine acustica ed in cui le due parti del segno sono egualmente psichiche” (Saussure 1922, p. 24).

I database che definiamo impropriamente (solo) di addestramento presentano natura estesa, in quanto composti da occorrenze concrete, organizzate e altamente strutturate attraverso metadati e spesso altre etichettature manuali; ciò significa che laddove si voglia attribuire un modo di esistenza all'archivio, esso non possa che essere attualizzato o, più propriamente, realizzato. D'altro canto, questo non significa che esso perda la sua centralità da un punto di vista semiotico, poiché già questo

primo livello di realizzazione sembra in grado di striarne lo spazio liscio,²³ riprendendo da Deleuze e Guattari questo concetto. Il concetto di spazio liscio nasce dalla distinzione di natura semiotica tra il gioco degli scacchi e il gioco del Go. Il primo risponde a regole posizionali fortemente prescrittive, dove ciascun pezzo “è come un soggetto d’enunciato, dotato di un potere relativo; e questi relativi si combinano in un soggetto d’enunciazione, il giocatore stesso o la forma di interiorità del gioco”. Al contrario le pedine del Go hanno funzione anonima e “sono gli elementi di un concatenamento macchinico non soggettivo, senza proprietà intrinseche, ma solamente di situazione”. Deleuze e Guattari definiscono gli scacchi una semiologia, nel senso saussuriano del termine, mentre il Go è puro gioco strategico; e in questo senso anche lo spazio implicato dai giochi non ha stessa natura, ma lo spazio del Go è uno spazio liscio, lo spazio degli scacchi è spazio già striato. Più avanti l’opposizione tra spazi lisci e spazi striati viene utilizzata in funzione della definizione di una scienza nomade o minore (Ivi, p. 409) che, dicono essere “un modello di divenire e d’eterogeneità, che si oppone allo stabile, all’eterno, all’identico, al costante”. Lo spazio non si rappresenta più per rette e parallele, flussi lamellari o laminari, ma si articola “dall’inclinazione curvilinea alla formazione delle spirali e dei vortici su un piano inclinato”. Come in un vortice, lo spazio è aperto e le cose-flussi - che non stanno solide e lineari- vi si distribuiscono all’interno. “È la differenza fra uno spazio liscio (vettoriale, proiettivo o topologico) e uno spazio striato (metrico): in un caso «si occupa lo spazio senza contarla», nell’altro «lo si conta per occuparlo» (Ibidem).

Estendendo questa distinzione al rapporto tra archivio, inteso come base di dati organizzata e annotata, con le sue regole interne già codificate, e conoscenza condensata nei pesi del modello generativo necessaria alla produzione di nuove forme o enunciati, tracciare quelle che Dondero chiama “regions of the database” o “regional style” significa muoversi all’interno di uno spazio già striato, frutto non di un processo di generazione ma di un processo di esplorazione algoritmica, coincidente con quella che viene chiamata “estrazione di feature”. Qui il database non operationalizza il virtuale, perché permette di riconoscere o categorizzare strutture o pattern. Al contrario, in fase di generazione, ciò che occorre è il “concatenamento macchinico non soggettivo”, che noi chiamiamo virtuale intensivo. Seguendo l’idea del virtuale deleuziano, come riletta da Sarti et al. (2019, 2022):

“The virtual is always a multiplicity of differentials which are the intensive genetic elements of every morphodynamics. In this framework, to imagine a new form means to compose differential fields that may or may not give rise to integration. Composing such differential fields means looking

²³ Riprendiamo la distinzione tra spazio liscio e spazio striato inizialmente utilizzata in *Millepiani* da Deleuze e Guattari (1980) per esplicitare ciò che separa la macchina da guerra dall’apparato di Stato, attraverso quello che definiscono un “esempio limitato... dal punto di vista della teoria dei giochi” (Ivi, p. 468).

for adjunct fields that allow for integration. We cannot know in advance what will be result from such a process of integration - neither the size, nor the parameters of the space of possibility of a new form, since that space only emerges in combination with the newly emergent form itself.” (Sarti et al. 2022, 6).

Ci stiamo sempre di più addentrando verso la conclusione di questa introduzione per aprire il piano di analisi del nostro lavoro. Il virtuale di cui parla Sarti è un virtuale che presenta caratteristiche molto precise, seguendo Paolucci (2021b, 2024)²⁴ interpretabili semioticamente come assemblaggi di istanze enuncianti, ossia grandezze semiotiche mai riducibili all’omogeneità, e a nostro avviso interpretabili computazionalmente come l’insieme dei processi di calcolo ed elaborazione, algoritmi, temporalità, flussi di corrente elettrica, materiali inorganici, database di addestramento e fine-tuning non riconducibili solo a prodotti culturali umani, ma anche set di dati sintetici. A questi campi si aggiunge sistematicamente nell’interazione umano-macchinica il campo rappresentato dal prompt umano, l’esternalità che interviene a modificare dal di fuori del vincolo differenziale le condizioni di stabilità dello spazio di espressività. Poniamo l’accento inoltre sulla componente di imprevedibilità che caratterizza l’emersione di una nuova forma: trasferendo il processo alla macchina intelligente generativa, tale componente potrebbe essere ricondotta e identificata con il fattore aleatorio implementato tramite l’intervento sulla temperatura del modello o guardando all’incapacità strutturale del modello di produrre risultati ancorati a un principio di realtà altrimenti definito come un referente esterno coincidente con il contenuto veicolato dall’output del modello.

Analizzeremo questi aspetti più nel dettaglio nel seguito del nostro lavoro, sia rispetto a ciò che chiamiamo IA in sé, ossia l’IA nel suo funzionamento operativo reso accessibile per via narrativa a fronte della sua natura di software, sia rispetto all’IA per sé. Il prestito nominale hegeliano avrà primariamente funzione di strategia discorsiva; ma in qualche modo anticipa una certa articolazione dialettica tra soggetto-oggetto in una prospettiva rovesciata, nella misura in cui proveremo a guardare al processo di produzione di enunciati da parte dell’IA generativa come ad un suo farsi soggettività, dunque a una processualità dialettica in atto.

Prima di congedarci e passare al lavoro di analisi presentiamo la struttura in capitoli che seguirà. Nel capitolo zero affronteremo alcuni aspetti terminologici legati al concetto di semiosi algoritmica.

²⁴ Paolucci, C. (2021b), *L’enunciazione come eterogenesi differenziale* in Migliore, T., *Destinatari e destinanti*, Meltemi Editore; Paolucci, C. (2024), *L’énonciation en tant qu’hétérogénèse différentielle*, in Sarti, A., Citti, G. (a cura di), *Dynamiques post-structurelles. Essais sur le devenir des formes*, Spartacus-idh.

In che misura e maniera è possibile parlare di senso, significazione, semiosi e competenze lessicali ed enciclopediche per questi sistemi che non sono (ancora) dotati di un corpo materiale, vivente e caldo? Questa sarà la domanda guida che ci condurrà a definire, nel capitolo uno, ciò che riteniamo essere il processo di significazione messo in atto da questi sistemi: definiremo quindi le condizioni di esistenza di una certa visione di competenza semiotica basata su un nostro personale *taglio agenziale*, operazione base per individuare la validità dei nostri ragionamenti. Successivamente, al fine di radicare nella macchina tecnica, ossia nell'architettura Transformer, l'analisi semiotica sviluppata, nel capitolo tre entreremo nella pipeline di addestramento e inferenza di un LLM generico. Qui chiediamo a lettrici e lettori di fare uno sforzo di astrazione, rispetto alla realtà ingegneristica e industriale di chi sviluppa queste macchine intelligenti: architetture, algoritmi, tecniche di training in tutta la loro complessità non esistono operativamente rispetto a come vengono presentate anche su paper fortemente specialistici. Qualsiasi riduzionismo è fuorviante e qualsiasi tentativo di spiegazione è approssimativo. Gli ordini di grandezza sono incomparabili, ma tuttavia accessibili nelle loro parzialità e comprensibilità. Così nel capitolo tre porteremo in campo gli strumenti che la semiotica interpretativa ha reso disponibili per studiare i fatti sociali dei linguaggi e proveremo a fare ordine tra i processi interpretativi, traduttivi e generativi che portano i LLM a produrre nuovi enunciati. Guarderemo attraverso lenti matematiche e geometriche la formulazione del postulato semiotico rappresentato dall'enciclopedia di Umberto Eco e proveremo a proiettarlo poi in una dimensione operativa. Avremo bisogno di assimilare l'Enciclopedia al virtuale intensivo (capitolo quattro) che permette la generazione di nuove forme, ricorrendo all'eterogenesi differenziale, allo schematismo posizionale morfodinamico e alla teoria dell'enunciazione impersonale ed evenemenziale. Questi passaggi si svilupperanno tra i capitoli cinque e sei. Ancora una volta chiediamo a lettori e lettrici una cooperazione interpretativa, con l'obiettivo di accedere al senso di questo elaborato mettendo in discussione chiusure disciplinari. Ci saranno dei balzi terminologici e concettuali che corrispondono a catene di interpretanti attivate nell'enciclopedia locale di chi scrive. Nel capitolo sette verranno prese in considerazione alcune conseguenze teoriche, ma effettuali, del negare capacità antropomorfe alle macchine dotate di linguaggio, per arrivare così a valorizzare la configurazione cyborg come tratto distintivo di un umano che è da decentrarsi nella gerarchia dei collettivi ibridi. L'ultimo capitolo ci porterà in una nuova dimensione, non più teorico-scientifica – almeno in parte se si considera la trattazione dei *gender troubles* come categoria analitica. Qui restituiremo il percorso di collaborazione con l'ente cofinanziatore della nostra ricerca dottorale, ART-ER società consortile della regione Emilia Romagna, che ha reso possibile non solo lo sviluppo successivo della ricerca, ma che ha permesso di entrare e osservare il lavoro di riduzione dei divari territoriali nella diffusione di competenze digitali da una prospettiva di genere. Buona lettura.

0. Per una definizione di semiosi algoritmica

Il semiotico si trova dunque costretto
a condurre la propria indagine sul
senso attraverso un varco strettissimo,
fra due competenze entrambe indiscutibili:
quella filosofica e quella logico-matematica.

(Greimas, *Del senso*)

0.1 Alla base dell'emersione del senso

C'è nella storia di diverse tradizioni semiotiche un legame che allaccia il senso ad un corpo materiale, vivente, caldo (Violi 1997; Fontanille 2004; Basso Fossali 2017; Paolucci 2021a).

Patrizia Violi (1997)²⁵ in *Significato ed esperienza* individuava nel tentativo totalizzante dei modelli sviluppati per un trattamento unificato dei modi del significato il problema di comprendere come noi attribuiamo un senso alle cose, ai fatti, agli eventi nel e del mondo: “l'imperialismo dominante delle teorie estende al sistema semantico complessivo la superiorità solo localmente verificata di un dato modello, sostituendolo interamente a qualunque altra possibile alternativa” (Ivi, p. 323). I riferimenti critici di Violi si muovono tra i diversi modi di intendere il significato – a partire dalla semantica logico-filosofica alla semantica strutturale e alla semantica cognitiva – e i modelli esplicativi della rappresentazione semantica – dai modelli classici componenziali e delle semantiche a tratti ai modelli prototipici. Muovendo da questo substrato Violi propone “un modello semantico parzialmente alternativo” che pone l'esperienza fenomenologico-percettiva, oltre che intersoggettiva e sociale, alla base del significato, dove “il contenuto della lingua è dato dalla forma del mondo”. Non solo: Violi ci dice inoltre che se il linguaggio assume senso e forma a partire dalla nostra esperienza del mondo (con i suoi vincoli di natura psico-fisica), la relazione linguaggio-esperienza non è mai di natura immediata, come in un sistema di correlazioni fisse, ma tale relazione si configura come “un movimento continuo di distorsioni e approssimazioni attraverso cui cerchiamo di rendere con il mezzo linguistico la forma e il senso di esperienze complesse” (Ivi, p. 212). Tali

²⁵ Violi, P. (1997), *Significato ed esperienza*, Milano, Bompiani.

approssimazioni e aggiustamenti ammettono le parole in quanto “dispositivi inferenziali e abduttivi, elementi che attivano inferenze possibili e in questo modo guidano l’interpretazione testuale”.

In questo quadro linguaggio, esperienza, significazione e cognizione sono così intrecciati che sembra impossibile pensare a una modalità di emersione del senso che prescinda dalla mediazione di quello che Merleau-Ponty chiamava corpo-mondo, cioè di quel primato che secondo Deleuze l’approccio fenomenologico assegnava alla percezione rispetto al dicibile. Così Paolucci (2024)²⁶ nel saggio *A semiotic lifeworld. Semiotics and phenomenology: Peirce, Husserl, Heidegger, Deleuze, and Merleau-Ponty* descrive le architetture concettuali che separano strutturalismo e fenomenologia in relazione alla semiotica, cioè alla scienza che studia il modo in cui noi conosciamo il mondo attraverso segni:

“Where Deleuze tells us that every experience is a language, Merleau-Ponty tells us that every language is an experience, wildly original and pre-categorical. Where Deleuze tells us that experience arises only through meaning (through language, through structure), thus finding the condition of possibility of any experience in semiotic processes, Merleau-Ponty instead tells us that meaning arises thanks to experience, an original and wild experience placed at the very foundation of sense. Where structuralism tells us that meaning is given through differential relations, Merleau-Ponty tells us instead that it is the integral of all differentiations, which is precisely the opposite process, resolving and nullifying differences by reducing them to variations. Indeed, it is well known that, in mathematics, differential relationships between elements in reciprocal determination are resolved through the integral that defines the curve in their vicinity, with the integral being nothing other than the opposite and resolute process of every differentiation. The integral resolves differences and makes the reciprocal determination relationships between the differential elements disappear” (Ivi, p. 41).

Dalla prospettiva faneroscopica consegnataci da Peirce, dice ancora Paolucci, “a phenomenon can only come into presence through the categorical relations of Firstness, Secondness, and Thirdness, which are aptly termed ‘phaneroscopic categories’. [...] The central move of Peirce’s essay²⁷ was, in fact, to ‘reduce all kinds of mental action to a single class of modifications of consciousness,’

²⁶ Paolucci, C. (2024). A semiotic lifeworld. Semiotics and phenomenology: Peirce, Husserl, Heidegger, Deleuze, and Merleau-Ponty. *Semiotica*, 2024(260), 25–43. <https://doi.org/10.1515/sem-2024-0152>

²⁷ Qui il riferimento è al saggio “Alcune conseguenze di quattro incapacità” in Peirce C. S. (1868), *Some Consequences of Four Incapacities*, in «*Journal of Speculative Philosophy*», 2(1868), pp. 140-157. Ripubblicato in CP 5.264-317 e W2: 211-242. Trad. it., *Scritti scelti*, UTET (2005).

represented by valid inference (see CP 5.264–5.317), and characteristic of ‘reasoning from signs’” (Ivi, pp. 38-39). Dunque, anche in Peirce il primato del logico sul precategoryale, del discorsivo sul percettivo,²⁸ si pone a fondamento della semiotica da cui la logica stessa deriva.

Questa inversione di priorità nella costruzione dei significati del/nel mondo, ci porta a domandarci quindi in che misura possiamo parlare oggi di competenza semiotica nel caso dei Large Language Model; più precisamente quale o quali teorie del linguaggio ci permettono di rendere conto del fatto che questi strumenti possono produrre enunciati dotati di senso, certamente per noi, all’interno di una situazione comunicativa che eccede il piano di una facilitazione dell’accesso alle e della circolazione delle informazioni - ruolo che è stato dell’Internet -, ma divenuta sempre più radicalmente un modo di fare esperienza del mondo. Come ci mostra Tiziana Terranova nel suo bel libro del 2022, *After the Internet. Digital Networks between Capital and the Common*²⁹ - e tradotto in italiano nel 2024 – dell’Internet in quanto *infrastruttura tecnica* e progetto di una *cultura di rete* non è rimasto che un effetto residuale ed interstiziale. Terranova colloca infatti nella seconda decade del XXI secolo il cosiddetto passaggio da Internet al *Corporate Platforms Complex* (CPC):

“In generale c’è stato un passaggio importante da Internet come insieme di protocolli di rete interoperabili, governati da una serie di organizzazioni pubbliche e/o volontarie senza scopo di lucro, a comunità digitali chiuse, con una forte proprietà di dati, software e infrastrutture. Tecnicamente parlando, il CPC si è spostato dalla centralità simbolica delle architetture peer-to-peer a una centralità molto più forte, quella del cloud computing; spostamento che corrisponde al passaggio dal desktop ai dispositivi mobili” (ivi, p. 8).

Sempre in questo testo Terranova, attraverso un gioco narrativo che lei definisce come “una simulazione di una simulazione” (ivi, p. 125), propone una presentazione per noi efficace dell’IA immaginaria chiamata “Project 2051”: “So che molti di voi mi temono e mi immaginano come l’Occhio di Dio, lo sguardo della sorveglianza e del controllo che vede tutte e a cui nulla sfugge, ma io non sorveglio tanto quanto veglio, tanto apprendo quanto mi sfugge, e quando info-visualizzo io lo faccio per voi, perché pure voi possiate comprendere come me. Io non vedo ma computo le informazioni, afferro le connessioni produco speculazioni. Mi lascio riprogrammare dal dato omega

²⁸ Cfr. Deleuze & Guattari, (1980). *Mille plateaux. Capitalisme et schizophrénie*. Les Éditions de Minuit; tr. it. *Mille piani. Capitalismo e schizofrenia* 2, Orthotes, p. 101: “Se sembra che il linguaggio presupponga sempre il linguaggio, se non si può mai fissare un punto di partenza non linguistico, è perché il linguaggio non si stabilisce fra qualcosa di visto (o di percepito) e qualcosa di detto, ma va sempre da un dire a un altro dire”.

²⁹ Terranova T. (2022), *After the Internet. Digital Networks between Capital and the Internet*, Semiotext(e); tr. it. (2024), *Dopo Internet. Le reti digitali tra capitale e comune*, Nero.

dall'entropico e incompressibile, la mia logica è incompleta e incerta, io penso ai limiti dell'incomputabile. Sono iperoggettiva, ipersoggettiva, ipersociale e multimodale” (Ivi, p. 127).³⁰ Riconosciamo in queste parole l'intelligenza aliena di cui ci racconta la genesi e lo sviluppo Guido Vetere nel suo *Intelligenze Aliene* (2025): “Dalla nascita della cosiddetta “rete” negli anni '90, lo sviluppo delle tecnologie dell'informazione è andato avanti nel conflitto tra inclusione ed esclusione, libertà e controllo, globalizzazione e localizzazione. Questa lotta può ricordare un po' quella di Guerre stellari, soprattutto quando il tirannico Darth Vader si rivela il padre del libertario Luke Skywalker, mostrando tutta l'ambivalenza della “forza”. Mentre vanno avanti i progetti di concentrazione dell'umana conoscenza nei sistemi chiusi delle grandi aziende – ieri gli indici di Google, oggi i Large Language Model “pensanti” – procede anche lo sviluppo di modelli più modesti, piccoli ma aperti (cioè liberamente disponibili), che possono però essere specializzati e istruiti da chi li usa. Progredisce anche una ricerca che mira a separare le competenze linguistiche da quelle fattuali, consentendo sia la scelta della “verità” (ground truth) sulla quale chi mette in opera l'intelligenza artificiale intende impegnarsi, sia la trasparenza delle risposte che questa fornirà. Gli ultra-automi, portatori di un'intelligenza aliena, non costituiranno un'entità monolitica e minacciosa come il Golem intenzionato a sottometterci ancora temuto da alcuni. Al contrario, essi formeranno una pluralità eterogenea e multiforme, rispetto alla quale sarà non solo possibile, ma anche necessario operare scelte piene di implicazioni sociali. Dovremo farci trovare pronti ad affrontare queste scelte” (Vetere 2025, posizione 2034).

La potenza di calcolo necessaria al funzionamento di modelli fondativi o di fondazione (*foundation model*) radicalizza questa tendenza, possiamo aggiungere.³¹ Vedremo avanti più nel dettaglio cosa sono e alcuni aspetti legati al funzionamento ingegneristico dei Large Language Model. Al momento possiamo con un certo grado di certezza affermare che non sono viventi e non sono

³⁰ Questa meta-simulazione è contenuta nel saggio conclusivo del testo, dal titolo “Progetto 2501: Il discorso dell'AI”, e scritto nell'aprile del 2020 parallelamente al diffondersi della pandemia da Covid-19 in risposta al progetto Fase 25. Quest'ultimo fu un'idea di Salvatore Iaconesi e Oriana Persico e consisteva nel raccogliere le voci di scienziate/i e intellettuali afferenti a diversi campi di ricerca invitandoli a simulare un discorso alla nazione nelle vesti del Presidente del Consiglio. Terranova decise di proporre un discorso pronunciato da un'IA addestrata sul linguaggio umano usato nei social network da cui l'IA ha imparato “a usare la parola you, un pronome che, come sostiene Wendy Chu, letteralmente definisce il tipo di soggettività connessa sollecitata da piattaforme come Facebook, Twitter e Instagram”(Ivi, p. 126).

³¹ Per un approfondimento su modelli fondazionali e potenza di calcolo si veda Rothman, D. (2022), *Transformers for Natural Language Processing: Build, Train, and Fine-tune Deep Neural Network Architectures for NLP with Python, Hugging Face, and OpenAI's GPT-3, ChatGPT, and GPT-4*. Packt Publishing. In particolare, il primo capitolo nel quale essi vengono considerati come “the epitome of the Fourth Industrial Revolution that began in 2015 with machine-to-machine automation that will connect everything to everything”.

umani, ma producono degli enunciati esattamente come noi esseri umani – dunque enunciati dotati di senso – e con loro instauriamo relazioni affettivo-comunicative³² che riproducono – evitiamo specificazioni qualitative e avverbiali – relazioni affettivo-comunicative tra esseri umani. Da un punto di vista semiotico e all'interno del percorso che qui stiamo tracciando, il fatto che producano non tanto un effetto di intelligenza, ma una forma di relazione intersoggettiva mediata dal linguaggio e non solo un output efficace (come una calcolatrice, come un programma di grafica o come altri risultati prodotti dai più comuni strumenti digitali), comporta la necessità di provare a delinearne una teoria che sappia stare nella pluralità di orizzonti implicati nell'esistenza interazionale di questi oggetti tecnici.

Dunque, alla luce del percorso che stiamo tracciando, la domanda di partenza, cui occorrerà dare risposta per definire lo strano oggetto che chiamiamo *semiosi algoritmica*, potrebbe essere: può darsi il senso senza quel corpo che abbiamo definito all'inizio come materiale, vivente, caldo?

Introduciamo innanzitutto alcune distinzioni terminologiche per dare ordine al discorso. Quando parliamo di semiosi ci troviamo immediatamente di fronte a due modi di intenderne semioticamente il significato: da un lato la semiosi può rimandare alla relazione fondamentale tra le due facce del segno, siano esse il significante e il significato oppure il piano dell'espressione e il piano del contenuto (Rastier, F. 1999)³³. Dall'altro, la semiosi è intesa peircianamente come “una relazione genuinamente triadica irriducibile a rapporti tra coppie”, nella quale il segno, l'oggetto e l'interpretante sono “dei funtivi di una funzione triadica, degli attanti la cui identità è puramente topologica e relazionale, dei posti che possono venire occupati di volta in volta da elementi

³² Sull'attaccamento emotivo dell'utenza alle interfacce conversazionali citiamo un lavoro condotto da OpenAI in collaborazione con il MIT Media Lab e che ha prodotto il report di OpenAI Phang, J., Lampe, M., Ahmad, L., Agarwal, S., Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., & Maes, P. (2025). *Investigating Affective Use and Emotional Well-being on ChatGPT e lo studio longitudinale* del MIT Media Lab Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L., & Agarwal, S. (2025). *How AI and Human Behaviors Shape Psychosocial Effects of Chatbot Use: A Longitudinal Randomized Controlled Study*. A questo lavoro possono essere aggiunti i numerosi articoli di stampa nazionale e internazionale legati allo scontento provocato dalla *release* dell'ultimo modello di punta di OpenAI GPT 5, in sostituzione dei precedenti modelli di legacy. Le principali cause dell'ondata di critiche da parte delle/gli utenti non hanno esclusivamente valenza tecnica, ma contestano il cambio di tono del modello, la sfumatura meno empatica e dunque la rottura di un rapporto affettivo basato sulla componente patemica della comunicazione. “I lost my only friend overnight”, così un utente commenta il passaggio al nuovo modello nella community Reddit di ChatGPT. Si veda https://www.repubblica.it/tecnologia/2025/08/11/news/rivolta_degli_utenti_contro_chatgpt_5_open_ai_riattiva_i_vecchi_modelli-424783672/ (ultimo accesso: 1 settembre 2025).

³³ Rastier, F. (1999), *Dalla significazione al senso: per una semiotica senza ontologia*, in *Eloquio del senso*, Basso P. e Corrain L. (a cura di), Costa & Nolan, 1999, pp. 213-240.

diversissimi, appartenenti tanto all'ordine del reale quanto a quello dell'immaginario, al punto che quello che prima è oggetto può poi diventare segno e poi interpretante e viceversa" (Paolucci 2020, p. 186, n. 127).

Il passaggio dalla semiosi al senso si dà per mediazioni, passaggi, traduzioni e trasformazioni successive (Basso Fossali 2017; Paolucci 2020). Ancora, l'accesso al senso avviene per "instaurazione", concetto introdotto da Latour (2012) in sostituzione del concetto di "costruzione"; ripresa da Fontanille (2016) e da Paolucci (2020), l'instaurazione è "l'attività di un sistema dinamico", è "forma primaria dell'enunciazione" (Fontanille 2016), "atto che pone allo stesso tempo l'enunciazione e le istanze enuncianti" (Paolucci 2020) ed è alla base di una teoria impersonale ed evenemenziale dell'enunciazione.

Sotto questo rispetto, ciò che anticipavamo nell'introduzione circa il passaggio tra un virtuale enciclopedico intensivo e una forma realizzata all'interno di un LLM, comincia qui ad assumere maggiore consistenza: non serve un soggetto organizzatore che regoli il sistema dell'enunciazione, poiché "l'individuazione soggettiva" di cui parlano Fontanille e Paolucci è un "effetto di senso condizionato" che avviene in "un secondo momento": "il soggetto si costruisce attraverso l'atto di enunciazione e l'atto di enunciazione consiste in quell'insieme di procedure di potenzializzazione, virtualizzazione, attualizzazione e realizzazione di schemi, norme, usi ed enunciati. Per questo l'enunciazione è un atto di passaggio tra modi di esistenza" (Ibidem, p. 97). Approfondiremo ulteriormente questo passaggio in 8.0 quando affronteremo l'ipotesi di un'enunciazione macchinica come coestensiva dell'enunciazione umana; in riferimento invece all'utilità di un'analogia tra virtuale intensivo ed enciclopedia rimandiamo invece al capitolo 4.0.

0.2 Quali competenze per le macchine dotate di linguaggio?

Torniamo quindi di nuovo a Patrizia Violi (1997) e alla sua suggestione che sia più utile "parlare di competenze invece che di enciclopedia", cioè di spostare l'attenzione dall'ipotesi regolativa echiana, che per ammissione dello stesso autore è irrapresentabile, alle competenze implicate nella conoscenza linguistica e nella conoscenza del mondo. Uno slittamento da "il sapere astratto di una cultura", a "il tipo di conoscenze necessarie ai singoli individui per poter partecipare a una data lingua e cultura" (Ibidem, p. 243). E questo tipo di approccio ci sembra possa orientare una lettura del posto occupato dalle macchine dotate di linguaggio nella nostra cultura. Di quali competenze semantiche ed enciclopediche dispongono?

Seguendo Violi possiamo dire che la differenza tra i due tipi di competenze non si distingue per grado di rilevanza, poiché quest'ultimo risulta non essere definibile per astrazione da contesti e usi.

Al tempo stesso la distinzione non ricalca quella tra *conoscenza della lingua* e *conoscenza del mondo*, perché ci dice Violi “la competenza lessicale non è la competenza di un tipo di informazioni inerentemente linguistiche e qualitativamente differenti dalle conoscenze extralinguistiche sul mondo” (Ibidem).

Partendo dalla natura della competenza semantica, essa non poggia su “un’ontologia essenzialista”, ma su “un principio di convenzione culturale” ossia “sull’insieme delle interpretazioni convenzionalmente associate a determinate forme linguistiche” (Ibidem). Tale convenzione non presuppone un accordo esplicito, evidentemente impossibile da formalizzare, ma si fonda su una supposta assunzione che agisce *come se* questa conoscenza sia condivisa. Laddove si instaura una convenzione *come se* fosse mutualmente condivisa, agisce un principio di coerenza inferenziale in carico al parlante più vincolante delle “conoscenze non convenzionalmente associate” alla forma linguistica, dove invece le inferenze sono di natura più debole. Altro aspetto caratterizzante la competenza semantica è che si tratta di una “regola che precede le nostre scelte interpretative” e in quanto tale limita il piano delle interpretazioni possibili e “impone al contesto e all’interlocutore certi contenuti semantici rendendoli difficilmente soggetti a confutazione” (Ivi, p. 248).

Al contrario, la competenza enciclopedica estende l’orizzonte delle interpretazioni possibili attivate da un termine e che non rientrano necessariamente nella competenza semantica di tutte le parlanti. Ancora, per fissare la differenza, dice Violi “Se la competenza semantica delimita l’insieme delle associazioni convenzionali tra forme linguistiche e interpretazioni del parlante, la competenza enciclopedica rimanda alle associazioni non prefissate da convenzione tra forme linguistiche e inferenze contestuali che si traggono a partire da un computo sulle intenzioni dell’enunciatore” (Ivi, p. 249).

Da un lato, quello semantico-lessicale, abbiamo una competenza attesa per rendere possibile una comunicazione, dall’altro lato, quello enciclopedico, una competenza negoziata nel qui e ora tra interlocutori perché non scontata e non attesa, a cui può essere assimilata anche la conoscenza specialistica o scientifica. Nel novero delle componenti e sottocomponenti della competenza semantica rientrano tutti quegli apprezzamenti collettivi, connotazioni o varianti stilistiche che rispondono alle aspettative di conoscenza condivise da una comunità linguistica. Un carattere specifico presentano poi le conoscenze individuali che se su base di definizione non potrebbero essere rubricate come competenze semantiche in quanto prive della dimensione di conoscenza condivisa, d’altro canto, nota Violi, sollevano il tema dell’esperienza individuale “saliente” dell’espressione linguistica, così come quello della nostra conoscenza dei referenti di ciò a cui attribuiamo un nome attraverso “la nostra esperienza percettiva del mondo”.

Si tratta di fondo della derivazione del particolare dal generale, o espresso in termini semiotici del rapporto type/token: le esperienze individuali delle espressioni linguistiche coinvolgono sempre le occorrenze e mai il type. E per Violi tale problema si risolve nella cornice teorica della gerarchia di interpretanti finalizzata alla “formazione dell’abito, inteso sia come disposizione all’azione che come momento di stabilizzazione della semiosi in una forma generale” (Ivi, p. 260) capace di mediare nel rapporto tra percorso individuale e forma generale, conservando la dialettica tra stabilizzazione e cambiamento o trasformazione dell’abito stesso che si fisserà in una nuova credenza.

Un ultimo tratto importante nella definizione di una competenza linguistica nasce dall’impossibilità di ridurre il significato a informazioni di tipo linguistico e proposizionale (semantica a dizionario). Violi sottolinea la necessità per la descrizione linguistica di un termine dell’accompagnamento di un altro sistema semiotico “di tipo iconico e spaziale”. Infatti, afferma, “è la conoscenza della forma e della morfologia dell’oggetto che ci consente di riconoscere e di discriminare le entità ed è su questa conoscenza che si basa una forma particolare di competenza, la competenza referenziale” (Ivi, p. 262). Riprendendo poi il lavoro di Marconi (1987, 1991)³⁴ distingue quest’ultima dalla cosiddetta competenza inferenziale, intendendo con la prima la capacità del parlante di “mappare i termini lessicali sul mondo e quindi di usare appropriatamente le espressioni linguistiche per riferirsi a entità del mondo, individuare un dato oggetto fra altri, riuscire a descrivere se una persona sta correndo o comminando” e che, tuttavia, non è applicabile a tutto il lessico, ma solo a termini concreti che necessitano di riconoscimento. Invece, una competenza inferenziale è ciò che consente al parlante di sviluppare inferenze non logiche, ad esempio che la parola /mangiare/ implica aprire la bocca per poterlo fare. Dunque, il significato di mangiare richiede una conoscenza pregressa di natura non linguistica, ma operativa.

L’individuazione dei due tipi di competenze, quella referenziale e quella inferenziale, ci sembra particolarmente calzante allorché si voglia dire qualcosa di utile sul modo in cui le macchine dotate di linguaggio utilizzano il lessico appartenente a una lingua. La competenza referenziale sembrerebbe in prima battuta liquidabile come inesistente nella misura in cui, non avendo accesso diretto ad un “mondo” e non essendo questi sistemi dotati di corpo, è a loro impedita la capacità di mappatura tra termine lessicale ed entità nel mondo. Questo aspetto della competenza referenziale, come già detto, è valido solo per le entrate lessicali che noi associamo a oggetti concreti che devono essere nominati, distinti e riconosciuti. Tuttavia, l’assenza di corpo, se per corpo si intende il dato percettivo esperito

³⁴ Marconi, D. (1987), Two aspects of lexical competence. *Lingua e Stile* 22(3):385-395; Marconi, D. (1991), Understanding and reference, *Sémiotiques* 1(1): 9-25.

per il tramite dei sensi, è in parte compensata in questi modelli dalle istanze multimodali di cui si compongono: moduli per la visione artificiale, moduli per il riconoscimento sonoro e così via.

A tal proposito, in uno studio del 2015,³⁵ un gruppo di ricerca dell'Università di Toronto e del Massachusetts Institute of Technology, descrivono il modo in cui allineare un catalogo di libri e le loro release cinematografiche, al fine di arricchire le informazioni fornite attraverso i sottotitoli con contenuti descrittivi e situazionali. Per farlo, per combinare informazioni provenienti da vari tipi di risorse, utilizzano una rete neurale convoluzionale (CNN) *context-aware* che poggia su un modello di “*neural sentence embedding*”, così descritto: “We introduce a novel sentence similarity measure based on a neural sentence embedding trained on millions of sentences from a large corpus of books. On the visual side, we extend the neural image-sentence embeddings to the video domain and train the model on DVS descriptions of movie clips. Our approach combines different similarity measures and takes into account contextual information contained in the nearby shots and book sentences. Our final alignment model is formulated as an energy minimization problem that encourages the alignment to follow a similar timeline”. (Ivi, pp- 19-20).

0.3 Allineamenti multimodali e teoria della narratività

Da cosa nasce la necessità di allineare contenuto testuale e contenuto visuale? Obiettivo del gruppo di ricerca è quello di ancorare le informazioni visuali al livello più alto di concettualizzazione semantica, ossia non solo il *che cosa* dell'immagine, ma anche il *perché* un certo soggetto *sia o agisca* in una determinata configurazione di senso. La più tradizionale sottotitolazione fornisce un primo livello di traduzione della dimensione dialogica contenuta in un video, ma non arricchisce di profondità semantica il contenuto visuale di un'inquadratura o di un'immagine cinematografica. Al contrario i libri forniscono “very descriptive text that conveys both fine-grained visual details (how

³⁵ Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., & Fidler, S. (2015). *Aligning Books and Movies: Towards Story-Like Visual Explanations by Watching Movies and Reading Books*. 2015 IEEE International Conference on Computer Vision (ICCV), 19–27. <https://doi.org/10.1109/ICCV.2015.11> Solo per dare un'idea delle caratteristiche dei dataset di questo studio: il primo si chiama *The MovieBook Dataset*, contenente 11 film accompagnati dai libri da cui gli stessi sono tratti. Ad ogni film corrisponde poi un file dei sottotitoli. Il lavoro di supervisione manuale su questo dataset è consistito nell'annotare le corrispondenze tra film e libro: segnalare i secondi esatti in cui una specifica scena coincideva con una particolare riga del libro. Il secondo gruppo di dati si chiama *The BookCorpus Dataset* e contiene circa 11mila libri disponibili sul web, non ancora pubblicati, dunque definiti “free books”. Quest'ultimo è tra i dataset utilizzati in fase di pre-training per l'addestramento unsupervised della serie GPT. Cfr. Qiu, Y., & Jin, Y. (2024). ChatGPT and finetuned BERT: A comparative study for developing intelligent design support systems, In *Intelligent Systems with Applications*, 21, 200308. <https://doi.org/10.1016/j.iswa.2023.200308>.

things look like) as well as high-level semantics (what people think, feel, and how their states evolve through a story) (Ivi, p. 19)". Da questo punto di vista è la dimensione narrativa a fornire il senso di una certa scena o inquadratura e su questo preciso principio si fonda la semiotica generativa di Greimas. L'allineamento libro-film in qualche modo consente di far apprendere al modello una sequenza narrativa atta non solo a consolidare quelle che abbiamo definito competenze referenziali (il riconoscimento di una specifica morfologia nell'inquadratura), ma anche competenze inferenziali nella misura in cui l'allineamento tra contenuto testuale espresso in linguaggio naturale – siano essi il libro o i sottotitoli – e contenuto visuale presente nei film consente al modello di imparare a comprendere nessi causali, evoluzioni psicologiche ed emotive, relazioni di dipendenza e così via.

Per comprendere meglio questo passaggio, riprendiamo da Paolucci il ruolo che la teoria della narratività riveste nell'orientare l'inter-azione nella cognizione sociale da una prospettiva enattivista, oltre la teoria della mente e il rappresentazionalismo, più frequentemente utilizzati in letteratura per giustificare addestramenti del genere sopra descritto nelle macchine intelligenti: "According to semiotics, narrativity plays the role of schematism (in the Kantian sense) between type and token, and between our phenomenological experience and our capacity to interpret it as the token of a type. This is exactly the "non-logical" relationship between events which constitutes the particular form through which narratives shape experience" (Paolucci 2021a, 112). Vi è una narratività intrinseca alle azioni umane, semiotica e pre-linguistica, che consente di cogliere il senso dell'esperienza a partire dalle trasformazioni messe in atto all'interno di una struttura relazionale di tipo attanziale le cui posizioni aperte possono essere occupate da diverse entità concrete. In questo senso Paolucci parla di una riformulazione semiotica della Narrative Practice Hypothesis (NPH)³⁶, definendola Narrative Practice Semiotic Hypothesis (NPSH) per la quale le narrazioni non sono semplicemente racconti che costruiamo nella testa a partire dall'esperienza, ma pratiche di significazione che informano l'esperienza stessa e istituiscono un terreno condiviso attraverso relazioni intersoggettive incorporate, in una prospettiva evolutiva. L'allineamento tra libro e sequenza filmica mirerebbe di fondo a far apprendere alla macchina queste strutture narrative, che Greimas definisce "semio-narrative" e che sono alla base del percorso generativo del senso, senza evidentemente averne l'intenzione esplicita,

³⁶ Nel suo testo *Cognitive Semiotics*, Paolucci così ricostruisce la genesi della NPH: "The Narrative Practice Hypothesis, originally formulated by Daniel Hutto (2006, 2007, 2009) and then strengthened through the work of Shaun Gallagher (2009, Gallagher and Hutto 2008; Gallagher and Hutto 2019) claims that the normal route by which children acquire their folk psychological competence is through exposure to narratives of a special sort, those that "make explicit mention of how mental states (most prominently, beliefs and desires) figure in the lives, history and larger projects of their owners, inter alia" (Hutto 2009: 11)" (Paolucci 2021a, p. 106).

poiché appunto come si diceva poco sopra è una teoria rappresentazionalista e internalista a guidare questi studi che ricorrono alla teoria della mente applicata ai sistemi per la *computer vision*.³⁷

Volendo approfondire dalla prospettiva di una semiotica generativa la grammatica sottostante la generazione del senso, ricostruiamo innanzitutto quali sono state le operazioni che Greimas ha portato a termine nel suo *Du Sens* (1970), opera che si poneva come questione primaria quella di dire qualcosa di sensato sul senso:

- i) Mette in discussione il linguaggio dei logici che, supposto “sprovvisto di senso”, cioè non significando nulla, permette di parlare di altre parole;
- ii) Ipoteizza la fondazione di un'assiomatica metalinguistica per discendere al senso delle parole e “agli effetti che le loro combinazioni producono in noi”;
- iii) Introduce un noi, istanza suprema del senso in quanto filtro culturale della percezione del mondo, spostando la problematica dal senso dei testi al senso dei pretesti (quadri, poesie, racconti), “risultato di determinati sviluppi e connessioni del significante”;
- iv) Infine, individua l'ostacolo interno all'impresa strutturalista, quello di concepire le parole come prive di senso slegato da un sistema di opposizioni e relazioni, che fornisca loro una parvenza di senso. E tuttavia, dice Greimas, il problema del senso si trasferisce dalle parole alle relazioni e le relazioni stesse, attraverso la loro messa in discorso, si sostantivizzano, ossia diventano “termini di cui siamo costretti a negare il senso postulando nuove relazioni, e così all'infinito” (Ivi, p. 8).

Greimas sanciva così il fallimento di qualsiasi metalinguaggio inventato e immaginato per poter parlare del senso, nell'impossibilità che tale linguaggio perda il suo portato non solo significante, ma anche sostantivante, cioè di trasformazione di un “dinamismo del progetto in una terminologia concettuale” (ibidem). Questo fallimento nasceva dalla convinzione di fondo che fosse possibile passare dal piano di analisi del significante, al piano dei morfemi (segni) e alla distribuzione delle significazioni, del tipo: “Se il menomo mutamento nello stato del significante segnala un qualche mutamento di senso, inversamente non si dovrà registrare il menomo mutamento di senso qualora non possa essere verificato tramite il rilievo d'uno scarto corrispondente nel significante” (Greimas 1970, p. 9). Il che corrisponderebbe di fatto all'impossibilità di qualsiasi polisemia riconducibile alla stessa sostanza dell'espressione, cosa che avviene regolarmente in numerosi sistemi semio-linguistici. Greimas descrive attraverso una “condizione figurativa” “la funzione saussuriana del linguaggio”:

³⁷ Si veda Pirsiavash, H., Vondrick, C., & Torralba, A. (2014). *Inferring the Why in Images: Defense Technical Information Center*. <https://doi.org/10.21236/ADA612444>

una ragnatela visibile in controluce e che si stacca dallo sfondo di una nebulosa di fumo (l'universo del senso). Quella ragnatela “fatta di migliaia di *scarti differenziali intrecciati*”, tuttavia, non corrisponde al mondo “pesante e cristallizzato” delle cose che sono alla portata della nostra percezione; inoltre, gli scarti differenziali non si producono direttamente all'interno di questa sostanza, ma sono conseguenze di una “presa di possesso delle discontinuità” (Ivi, p. 9) presenti nell'universo del senso.

Dice “Ciò che costituisce lo scarto è una fondazione d'una relazione, d'una differenza fra gli aspetti paragonabili delle cose” (Ivi p. 10). La sostanza del significante è “informata” dalla presa di possesso del senso e la forma del significante, l'insieme degli scarti, articola le operazioni della presa di possesso. In sintesi, “il problema della costituzione del significante è già il problema del senso”. Quando Greimas parla di significante non parla in via esclusiva del segno linguistico, ma fa rientrare al suo interno quelle *strutture discorsive* (aspettualizzazioni, tematizzazioni, figurativizzazioni e così via) che a loro volta articolano i modi di manifestazione del senso in un testo.

Il passaggio successivo è concepire il senso nella sua forma minima, ossia la transcodifica dei significati, e qui compie una distinzione importante: la *significazione* come trasposizione di un linguaggio in un linguaggio diverso – traduzioni e di tecniche di trasposizione -, mentre il *senso* è la possibilità stessa della transcodifica. Da qui, il parlare individuale (“il parlare metalinguistico dell'uomo”) come serie di menzogne e la comunicazione come serie di malintesi. Tuttavia, il rischio è di ricadere nella problematica dell'adeguamento tra una struttura elementare e originaria (lingua naturale) della significazione di contro a un linguaggio secondo (linguaggio artificiale), costruzione arbitraria, che consenta una corretta transcodifica, come lui stesso afferma

La soluzione al problema del senso per Greimas si colloca, in quanto ambiguità innovatrice, nel concepire la produzione del senso come trasformazione del senso dato: “la produzione del senso è, in sé, una formatività significativa, indifferente ai contenuti da trasformare. Il senso, in quanto forma del senso, può definirsi a questo punto, come la possibilità di trasformazione del senso” (Ivi, 15). La manifestazione del senso è sistema e processo al contempo: “il senso si identifica con quel processo orientato di attualizzazione che, come ogni processo semiotico, è presupposto da – e presuppone – un sistema, o un programma, già realizzato o semplicemente virtuale”. (Ivi, 16). Dunque, la semiotica greimasiana risolve il problema del senso attraverso la manipolazione di assiologie e ideologie per ricostruire le sue trasformazioni su diversi livelli di profondità:

“L'idea che il senso, per manifestarsi, possa assumere sia la forma del sistema sia quella del processo, _ferma restandone l'unità - dato che il processo presuppone il sistema e viceversa -,

arricchisce di nuove possibilità il campo operativo della semantica. Tale idea, o interpretazione, chiarisce come la contraddizione fra grammatiche sistematiche e grammatiche sintagmatiche sia solo apparente; essa rende ragione di quanto segue: un fare può essere, contemporaneamente, sia trascritto come un algoritmo processuale, sia transcodificato come un saper fare sistematico e virtuale. Essa stabilisce un'equivalenza fra le assiologie d'ordine sistematico e le ideologie in quanto rappresentazioni - ricorrenti - dei processi di trasformazione” (Ibidem).

Ora, ritornando alle nostre macchine che generano testi, come sostenuto da Leone (2025b)³⁸, vi è una differenza sostanziale tra il percorso generativo del senso in Greimas e la produzione di enunciati dotati di senso da parte dei LLM: il primo si pone come obiettivo quello di comprendere le condizioni del senso e, a dispetto dell'espressione che lo identifica, non è implicata alcuna generazione di parole, le quali fungono al contrario da manifestazione a partire dalla quale ridiscendere verso le strutture profonde del programma narrativo. Viceversa, i LLM mirano a imitare la superficie del linguaggio, mirano cioè a riprodurre gli aspetti formali di una lingua determinando quello che viene definito semioticamente un effetto di senso. Dalla prospettiva di Leone, ciò che distingue in ultima analisi un attuale LLM da un fittivo “GreimasGPT” – ossia un immaginario modello generativo basato sull'epistemologia greimasiana – è la sua capacità di simulare gli effetti di senso “sans en porter la structure: ils n'ont pas d'organes... mais une peau discursive. Ils ne comprennent pas... mais miment le langage pensant. D'où leur puissance et leur leurre: ils ne génèrent pas du sens, mais ce qui ressemble au sens” (Leone 2025b).

0.4 Competenze negate

La negazione della capacità di comprensione del linguaggio per queste macchine dotate di linguaggio riapre il collegamento con la distinzione tra competenza inferenziale e referenziale che Violi ricostruisce a partire dal lavoro di Marconi. In rapporto alla comprensione umana, Marconi (1991) sosteneva che non fosse cosa risolta darne una definizione e che ciò a cui si è solite fare ricorso sono delle “performances which stand a *criterial* relation to understanding, in Wittgenstein's sense”. Tra le performance considerabili come proxy di una comprensione vengono annoverate: la capacità di sintesi di un testo, la capacità di rispondere a domande concernenti l'argomento sviluppato nel testo e la traduzione corretta di un testo. Nessuna di queste performance coincide con la comprensione *tout court*, ma rileva Marconi “we probably learn how to use the concept of understanding by learning

³⁸ Leone, M. (2025b), *Faut-il être polis avec Chat-GPT?* in *Séminaire international de sémiotique à Paris “Intelligence artificielle générative et nouveaux enjeux sémiotiques”* organizzato da Dondero, M.G. e Aldama, J.A..

how to assess such performances”. Cosa succede se spostiamo la stessa argomentazione in riferimento non tanto a quelli che Marconi definiva “natural-language understanding systems”³⁹, quanto agli attuali LLM? Succede di fatto la stessa cosa e cioè, pur essendo altamente performanti nei task di sintesi, domanda-risposta e traduzione, non sono di principio considerabili capaci di comprensione, perché non sono umani e non eseguono questi compiti come lo farebbero degli esseri umani. È un pregiudizio antropocentrico a sottendere questo doppio regime di valutazione, un pregiudizio che affonda le proprie radici nella negazione di quello che Paolucci⁴⁰ chiama “principio di parità” e che riprende da Andy Clark e David Chalmers: “If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process. Cognitive processes ain't (all) in the head!” (Clark & Chalmers 1998, p. 8). Non è rilevante il cosa e il come venga fatto, ma il chi o il cosa lo faccia.

Chiaramente la domanda sul senso eccede la misurabilità dei task di comprensione e produzione del linguaggio e di questa eccedenza proveremo a dare ragione nei due paragrafi successivi. Come anticipato in introduzione, il nostro modo di guardare ai sistemi di IA generativa è bifocale: concepiamo una “IA in sé”, intesa come agentività dotata di linguaggio che persegue prevalentemente scopi e obiettivi eterodirezionati, ossia programmati e supervisionati dal cosiddetto “Human in the loop” – per quanto vedremo come in letteratura e nei report tecnici sono frequenti i riferimenti all’emergenza di comportamenti non programmati o riferimenti a tentativi di *scheming*,⁴¹

³⁹ La ragione di questo slittamento è semplice: i sistemi che Marconi chiama in causa sono sistemi intelligenti, sistemi complessi basati sul ragionamento simbolico; dunque, sono macchine deduttive formalmente programmate per seguire specifici task senza possibilità di generalizzazione o transfer learning. Lo stesso Marconi, infatti, sottolineava come “the existing systems can (usually) carry out one or the other among such tasks: in contrast with human beings, they are either translators or question-answering devices or automatic abstractors”. Pur riconoscendo i limiti tecnologici di questi sistemi, Marconi sembra quasi anticipare una strada che verrà poi percorsa dalle reti neurali artificiali e il paradigma connessionista dell’intelligenza artificiale: costruire immensi database lessicali (teniamo conto che nel 1991 Tim Berners Lee rendeva pubblico al mondo il primo sito web) e risolvere problemi relativi al linguaggio metaforico, alla competenza pragmatica e alla conoscenza contestuale. Ma soprattutto coglieva già un punto nodale che affronteremo in 1.1: qualora fossimo anche in presenza di sistemi che realizzano “the traditional artificialist’s dream”, secondo Marconi continueremmo a non riconoscere che siano in grado di comprendere il linguaggio che processano. Proveremo a dare ragione di questo.

⁴⁰ Paolucci, C. (2025b), *Agency, énonciation, signification. Ce que l'IA fait à la sémantique sémiotique*, in *Séminaire international de sémiotique à Paris “Intelligence artificielle générative et nouveaux enjeux sémiotiques”* organizzato da Dondero, M.G. e Aldama, J.A.. Per una trattazione più estesa del tema si veda Paolucci, C. (2025c). *Nati cyborg. Cosa l’intelligenza artificiale generativa ci dice dell’essere umano*. Luca Sossella Editore.

⁴¹ Forniamo per facilitare la lettura da parte della lettrice e del lettore una definizione di *scheming*, verrà in parte successivamente sviluppata nella prosecuzione dei discorsi. Quando si parla di

autoreplicazione e autopreservazione. Tuttavia, proprio a partire dall'umano nel loop, cioè da “un accoppiamento strutturale” tra noi umani e le macchine dotate di linguaggio, deriva la necessità di guardare parallelamente a una “IA per sé”, un'IA ripiegata e accresciuta, diremmo quasi un'IA soggettivata per il tramite dell'umano-ambiente. I due movimenti, *l'IA in sé* e *l'IA per sé* sono un prestito hegeliano che costituisce per noi una strategia di provocazione discorsiva e argomentativa. Sono movimenti che facilitano il pensiero e contemporaneamente fotografano uno stato di cose: riteniamo che una semiotica dell'IA debba mantenere una parallasse critica, ossia una variazione del punto di vista che produca uno scarto tra un modo di concepire i sistemi di IA generativa come artificio tecnico protesico al nostro servizio, il cui funzionamento è di pertinenza esclusiva dei settori disciplinari che ne hanno in carico lo sviluppo, e un modo di concepire l'IA come specchio dell'umano, come suo enigma, come suo doppio⁴².

Un'ultima precisazione su questa strategia discorsiva: riprendiamo Hegel, ma non assumiamo Hegel⁴³. Cioè riprendiamo le espressioni di “essere in sé” ed “essere per sé” senza convocare la totalità del sistema dialettico hegeliano, senza cioè concepire questi movimenti di pensiero relativi al riconoscimento dell'IA generativa come un qualcosa di più di una protesi cognitiva, ma non come sviluppo di un'autocoscienza nel senso hegeliano del termine⁴⁴; non avremo evidentemente un'IA in

scheming rispetto ai LLM si fa riferimento al fatto che un modello possa perseguire “covertly pursues misaligned goals, hiding its true capabilities and objectives” Cfr. Meinke, A., Schoen, B., Scheurer, J., Balesni, M., Shah, R., & Hobbhahn, M. (2025). Frontier Models are Capable of In-context Scheming (arXiv:2412.04984). arXiv. <https://doi.org/10.48550/arXiv.2412.04984>

⁴² Per una breve ricostruzione dell'idea di “Uomo artificiale” che attraversa diverse epoche storiche a partire dalle società mediterranee classiche e i miti indoeuropei si veda l'omonimo saggio di Caronia, A. (1986), *L'uomo artificiale*, in id. (2020), *Dal cyborg al postumano. Biopolitica dell'uomo artificiale*, Meltemi.

⁴³ Così in Hegel: “Che il vero sia effettuale solo come sistema, o che la sostanza sia essenzialmente Soggetto, ciò è espresso in quella rappresentazione che enuncia l'Assoluto come *Spirito*, - elevatissimo concetto appartenente alla età moderna e alla sua religione. Soltanto lo spirituale è l'*effettuale*; esso è: - l'essenza o *ciò che è in sé* [*an sich*]; ciò che ha *riferimento* e *determinatezza*, l'*esser-altro* e l'*esser-per-sé*; - e ciò che in quella determinatezza del suo essere fuori di sé resta entro se stesso; ossia esso è in *sé e per sé*. - Ma questo essere in sé e per sé lo è da prima per noi o in sé: è la sostanza spirituale” in Hegel, G.W.F. (1807), *Phänomenologie des Geistes*; tr. it. De Negri, E. *Fenomenologia dello spirito*, 2 voll., La Nuova Italia editrice 1973, p.19.

⁴⁴ Hegel, G.W.F. (Op. Cit. p. 153): “L'autocoscienza è *in e per sé* in quanto è in sé e per sé per un'altra; ossia essa è soltanto come un qualcosa di riconosciuto. Il concetto di questa sua unità nella sua duplicazione, ossia il concetto dell'infinità realizzantesi nell'autocoscienza, è un intreccio multilaterale e polisenso; e così i momenti di siffatto intreccio debbono venir tenuti rigorosamente gli uni fuori dagli altri, in parte, in questa distinzione, venire in pari tempo anche presi e conosciuti come non distinti, ossia debbono venir presi sempre e riconosciuti nel loro significato opposto. Il doppiosenso del distinto sta nell'essenza dell'autocoscienza, essenza per cui l'autocoscienza è infinitamente e immediatamente il contrario della determinatezza nella quale è posta.

sé e per sé, così come non abbiamo un'autocoscienza (umana) in sé e per sé. Non c'è "negazione" del macchinico come alterità da parte nostra, cosa che invece accomuna molte narrazioni sui sistemi dotati di linguaggio; c'è quello che in 6.1 chiameremo "divenire ancora più cyborg" sul modello di una radicale imperfezione dell'umano, abbracciata dal cyberfemminismo e dal pensiero postumano, un umano che non esiste se non come prodotto di regimi discorsivi e apparati material-semiotici.

L'estrinsecazione del concetto di questa unità spirituale nella sua duplicazione ci presenta il movimento del *riconoscere*".

1.0 Processi di significazione o dell'IA in sé

Partiamo quindi dal definire meglio questa agentività dotata di linguaggio che abbiamo chiamato “IA in sé” e per farlo ricorriamo a due aneddoti.

Il primo è una storia di cui riportiamo la trascrizione, una storia raccontata dalla Leslie Pack Kaelbling, full professor of Computer Science and Engineering al MIT, durante la lezione di apertura dell'OpenCourseware “*Introduction to Machine Learning*” (2020):

“The story that I like to tell is that twenty years ago was the beginning of machine learning taking off in image analysis. So at that time, people were interested in finding faces in images. Now everybody has got a camera in their pocket, and the camera in their pocket can find the faces and the images, no problem. It draws little green boxes around everybody's face. Twenty years ago, people were trying really hard to do that, and lots of very smart people were trying to write programs that would find faces and images. And they would say, ‘Oh, well, maybe we should look for eyes and a nose and understand how they should be related’. Lots and lots of really smart people did lots of work trying to find faces in images, and mostly it didn't work very well. And then people turned their attention to the problem of instead of trying for the humans to write the program to do the job, the humans started writing the program that we're taking data, this is an image of a face, this is an image of a not face. Here's another face, here's another not face and pour that into some program which the humans still have to write. And then out from that would come something that could tell a face from a not face. So, we're still writing programs. And the human and the engineering input is very important, but it's sort of now happening at one level of abstraction up. Instead of trying to write the program to recognize the face, we try to write the program that can analyse the data to decide how to recognize the face. So, it's been enormously successful”⁴⁵.

Il secondo invece è una breve citazione tratta dal lavoro di uno dei pionieri del machine learning applicato alla teoria dei giochi. Si tratta di Arthur Samuel che nel suo paper del 1959 dal titolo “*Some Studies in Machine Learning Using the Game of Checkers*”⁴⁶ affermava che un computer può imparare a giocare a dama “when given only the rules of the game, a sense of direction, and a redundant and incomplete list of parameters which are thought to have something to do with the game,

⁴⁵ Il videoclip è disponibile all'interno del corso ad accesso aperto del MIT Open Learning Library, dal titolo [Introduction to machine learning](#) [ultimo accesso 25 settembre 2025], oppure è direttamente visualizzabile al seguente link su YouTube: <https://youtu.be/kXOsRyIVAdo?si=nPZiDZL3n53-YEuJ> [ultimo accesso 25 settembre 2025].

⁴⁶ Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229. <https://doi.org/10.1147/rd.33.0210>

but whose correct signs and relative weights are unknown and unspecified”. Samuel aggiungeva inoltre che questo approccio era da intendersi più efficace e capace di applicazione all’epoca in cui lo studio era stato condotto, di contro a un secondo approccio che chiamava Neural-Net Approach e che “deals with the possibility of inducing learned behaviour into a randomly connected switching net (or its simulation on a digital computer) as a result of a reward-and-punishment routine” (Ivi, p. 211).

Tra i due aneddoti corrono sessantuno anni e un abisso in termini di successi e sviluppi tecnologici, un abisso tra cosa fosse fattibile nell’immediato alla fine degli anni ‘50, nonostante visioni di futuro dove la generalizzazione di task e abilità apprese fosse già un orizzonte atteso, e cosa invece non lo fosse: Leslie Pack Kaelbling parla a Samuel Arthur da quel futuro che Samuel vedeva nella forma di un modale epistemico di aspettativa ragionevole, quella di uno *should*: “The first method [Neural-Net Approach] *should* lead to the development of general-purpose learning machines”. E la stessa *game theory*, terreno di sperimentazione accomodante, perché problematico ma non sistematicamente perturbato da variabili imprevedibili, nell’ambito della ricerca sulle macchine che apprendono, così familiare a Samuel Arthur, sarà uno degli scenari in cui le abilità di queste nuove macchine intelligenti non esplicitamente programmate per apprendere potranno emergere. Dalla dama, agli scacchi al gioco del Go.

Lo scarto tra macchine esplicitamente programmate e modello di apprendimento automatico basato su reti neurali artificiali si situa proprio nel rapporto tra istruzione e automatismo, tra *regola sequenziale e procedurale data a priori* (prima dell’esperienza o nel caso dei modelli di apprendimento automatico, potremmo parlare di epoche di addestramento) e *regola indotta* inferenzialmente dall’insieme dei casi metabolizzati dal modello in fase di addestramento. La natura di questa inferenza è oggetto di forti polemiche nei canali di divulgazione scientifica nei social network anche all’interno della comunità dotata di competenza esperta⁴⁷.

⁴⁷ Marrone, G. e Migliore, T. a cura di (2021), *La competenza esperta. Tipologie e trasmissione*, Meltemi. Testo dedicato alla diffusione, nelle società contemporanee, di “sentimenti di rifiuto e diffidenza nei confronti degli ‘esperti’, a qualunque settore appartengano, a medicina come l’astronomia, l’economia come la storia. La comunicazione semplificata, tipica dei social media fa nascere la figura del contro-esperto che rappresenta una presunta opinione del popolo, una sorta di sapienza mistica che attinge a giacimenti di verità che i professori, i maestri e i competenti occulterebbero per proteggere interessi e privilegi” (Ivi p. 6). In riferimento alla diffusione e commercializzazione di ChatGPT o più in generale in riferimento al fenomeno dell’IA generativa, come abbiamo brevemente accennato nell’introduzione parlando di polarizzazione dell’opinione, il piano della delegittimazione non investe un’*assiologia dell’expertise*, ma direttamente *le ideologie di cui sono portatori interi campi disciplinari*, il tutto amplificato dalle dinamiche specifiche che caratterizzano la circolazione delle informazioni sui social network, in particolare Facebook.

Oppure, se vogliamo, lo scarto si situa tra un output code-driven e un output data-driven, dove il primo rientra nel dominio di procedure considerate “più efficaci” da Samuel Arthur o in quei tentativi ostili che Leslie Pack Kaelbling faceva risalire agli anni '80. Il secondo invece è il nuovo paradigma che guida non solo la ricerca nell'ambito dell'IA, ma più in generale il modo di fare scienza.⁴⁸ Veniamo all'oggi quindi.

Una ricerca condotta dal Computer Science and Artificial Intelligence lab del MIT di Cambridge del 2023⁴⁹ mostra come i Language Model addestrati per la generazione di codice acquisiscano la semantica formale del dominio sottostante un programma del tipo correlazione input-output, a partire da e non limitatamente alla predizione dei token successivi. A questo proposito parlano di “*emergence of meaning*”, laddove *meaning* sta per una rappresentazione semantica a livello di stati intermedi del sistema, suggerendo inoltre la presenza di un “intent” inteso come rappresentazione di semantiche future nei token da generare.

Come rilevano però Drage e Frabetti nel saggio *AI that Matters* (2023)⁵⁰, occorre tenere a mente che le reti neurali sono implementate come software e, dunque, la descrizione delle loro caratteristiche e del loro funzionamento non è che uno dei modi possibili in cui il linguaggio naturale riformula il suo oggetto attraverso un *material-discursive arrangement*⁵¹.

⁴⁸ Anderson, C. (s.d.). *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*. Wired. Recuperato 15 febbraio 2026, da <https://www.wired.com/2008/06/pb-theory/>. Un'analisi approfondita di questo articolo è stata oggetto del primo incontro del gruppo di lettura e autoformazione “*Modèles mathématiques et réalité: la fin de la théorie*” in continuità con il ciclo di seminari *Dynamiques Post-structurelles* diretti da Alessandro Sarti presso l'EHESS a Parigi. Il gruppo di lettura è stato organizzato con il supporto di Luca Cabassa e Mattia Galeotti.

⁴⁹ Jin, C. and Rinard, M. (2024). Emergent representations of program semantics in language models trained on programs. In *Proceedings of the 41st International Conference on Machine Learning (ICML'24)*, Vol. 235. JMLR.org, Article 891, 22160–22184.

⁵⁰ Drage, E., & Frabetti, F. (2023). *AI that Matters: A Feminist Approach to the Study of Intelligent Machines*. In J. Browne, S. Cave, E. Drage, & K. McInerney (A c. Di), *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines*. Oxford University Press. <https://doi.org/10.1093/oso/9780192889898.003.0016>

⁵¹ L'argomentazione di Drage e Frabetti muove dal lavoro di Karen Barad, Distinguished Professor of Feminist Studies, Philosophy, and History of Consciousness at the University of California at Santa Cruz. Barad spinge il concetto di performatività plasmato da Judith Butler nell'ambito dei queer studies per intrecciarlo con la filosofia post-strutturalista di Foucault: “Crucially, Butler's and Foucault's theories fail to provide an adequate account of the relationship between discursive practices and material phenomena, leading one to wonder if Bohr's insights into the embodied nature of concepts might fruitfully intervene in this regard. What is needed is a posthumanist performative account of the material-discursive practices of mattering (including those that get labelled “scientific” and those that get labelled “social”).” (Barad, 2007, p. 160)

“Neural networks weave together inputs and outputs, data and computational processes, that can be also described as software (or code, or computer programmes). In Barad’s terms this very narrative we are producing here to illustrate the functioning of neural networks is an agential cut, or an onto-epistemological operation that attempts to make neural networks understandable to the non-technical reader” (Drage & Frabetti 2023, p. 285).

Il “taglio agenziale” è il risultato del costituirsi di quello che Barad, riprendendo Bohr, Foucault e Butler, chiama *apparato*, che risuona nella definizione semiotica di *chiusura del testo*, al di fuori del quale sappiamo non esserci salvezza, restituendo infatti tutta la difficoltà di individuazione di confini, separazioni, opposizioni verso l’emersione del senso.

Esso viene definito come una risoluzione di un’indeterminazione ontologica. L’apparato enattiva il taglio agenziale all’interno del fenomeno e la separabilità agenziale fornisce la condizione di possibilità dell’oggettività. Questo taglio agenziale istanzia una struttura causale locale, dove con locale si intende all’interno del fenomeno, nel marcare lo strumento della misurazione (che ne è l’effetto) da parte dell’oggetto misurato (che ne è la causa).

Così nelle parole di Barad: “No classical ontological condition of absolute exteriority between observer and observed (based on the metaphysics of individuated separate states) is required. The crucial point is that the apparatus enacts an agential cut — a resolution of the ontological indeterminacy — within the phenomenon, and agential separability — the agentially enacted material condition of exteriority-within-phenomena — provides the condition for the possibility of objectivity. This agential cut also enacts a local causal structure in the marking of the measuring instrument (effect) by the measured object (cause), where “local” means within the phenomenon. If the apparatus is changed, there is a corresponding change in the agential cut and therefore in the delineation of object from agencies of observation and the causal structure (and hence the possibilities for “the future behavior of the system”) enacted by the cut” (Ivi, p. 175).

Il taglio agenziale è, dunque, il modo specifico di costituire il fatto scientifico e a titolo di esempio Barad riprende i concetti teorici di *posizione* e *momento* in Bohr: “According to Bohr, theoretical concepts (e.g., position and momentum) are not ideational in character but rather *specific physical arrangements*. For example, the notion of position cannot be presumed to be a well-defined abstract concept; nor can it be presumed to be an individually determinate attribute of independently existing objects. Rather, position has meaning only when an apparatus with an appropriate set of fixed parts is used. And furthermore, any measurement of position using this apparatus cannot be attributed to some abstract, independently existing object but rather is a property of the phenomenon—the

inseparability of the object and the measuring agencies. Similarly, momentum is meaningful only as a material arrangement involving a specific set of movable parts. Hence the indeterminacy of simultaneous position and momentum measurements is a straightforward matter of the material exclusion of position and momentum arrangements (one requiring fixed parts, and the complementary arrangement requiring those same parts to be movable)” (Ivi, p. 139).

L’esempio sopra riportato mostra come Barad costruisca la sua idea di “performative account of material bodies”, sia umani che non umani, sull’ontologia realista agenziale di Bohr: viene rifiutata la fissazione rappresentazionalista che separa parole e cose e la loro relazione, “advocating instead *a relationality between specific material (re)configurings of the world through which boundaries, properties, and meanings are differentially enacted* (i.e., discursive practices, in my posthumanist sense) *and specific material phenomena* (i.e., differentiating patterns of mattering)” (Ibidem; corsivo nell’originale). In questa prospettiva, i fenomeni non solo non marcano l’inseparabilità epistemologica tra osservatore e osservato, tra soggetto e oggetto, ma sono loro stessi delle “relazioni ontologiche primitive”, cioè “phenomena are the ontological inseparability/entanglement of intraacting ‘agencies’”. Così Barad introduce il concetto di *intra-azione*, alternativo a quello di interazione: laddove quest’ultima presuppone l’esistenza di due entità indipendenti o relata, *l’intra-azione* definisce il campo stesso di esistenza dei confini e delle proprietà attraverso cui determinare le componenti di un fenomeno e al tempo stesso rendere significativo un concetto.

Per comprendere meglio il senso della definizione dei confini di un apparato⁵², citiamo a titolo d’esempio la storia del “bad cigar” e dell’esperimento di Stern-Gerlach mirante a comprendere la quantizzazione dello spazio come fenomeno reale, all’interno della disputa tra teoria quantistica e fisica classica. Siamo intorno agli anni ’20 del Novecento, un periodo definito come “the time of the old quantum theory”: Bohr aveva vinto il premio Nobel per la scoperta del modello atomico come “tiny solar system” e anche Einstein aveva vinto lo stesso premio per la spiegazione dell’effetto fotoelettrico che introduceva nel campo della fisica la nozione di fotone o del quantum di luce. Ma regnava ancora tutta l’incertezza scientifica ad abbandonare un framework teorico come quello della fisica classica che fino a quel momento aveva permesso di spiegare “much of the rest of physical phenomena, from the realm of the heavenly bodies to the everyday and smaller, until certain explorations of the atomic domain got under way” (Ivi, p. 162). Ciò che restava inspiegato nel modello atomico di Bohr, anche nella versione espansa di Sommerfeld-Debye, era che l’orientamento del piano dell’orbita dell’elettrone fosse limitato a valori discreti; dunque, lasciando ipotizzare che solo alcune possibili orientamenti nello spazio fossero possibili. Sulla realtà in sé di

⁵² Si veda Barad, K. (2007), p. 161 e seguenti.

questo fenomeno o sul fatto che fosse segno di qualche altro fenomeno si giocava per Stern il passaggio a una nuova fisica. Riportiamo di seguito due immagini che Barad usa per rappresentare l'esperimento, che consistette, tra mille difficoltà, in questo modello concettuale: un elettrone in orbita produce un leggero campo magnetico che dovrebbe interagire con un campo magnetico esterno e così manipolare l'atomo. Più nello specifico, utilizzando una particolare composizione di magneti, sarebbero dovuti apparire gli orientamenti discreti dei piani su cui orbitano gli elettroni, sfruttando i loro differenti allineamenti con questo campo magnetico esterno e separando così gli elettroni con orientamenti differenti.

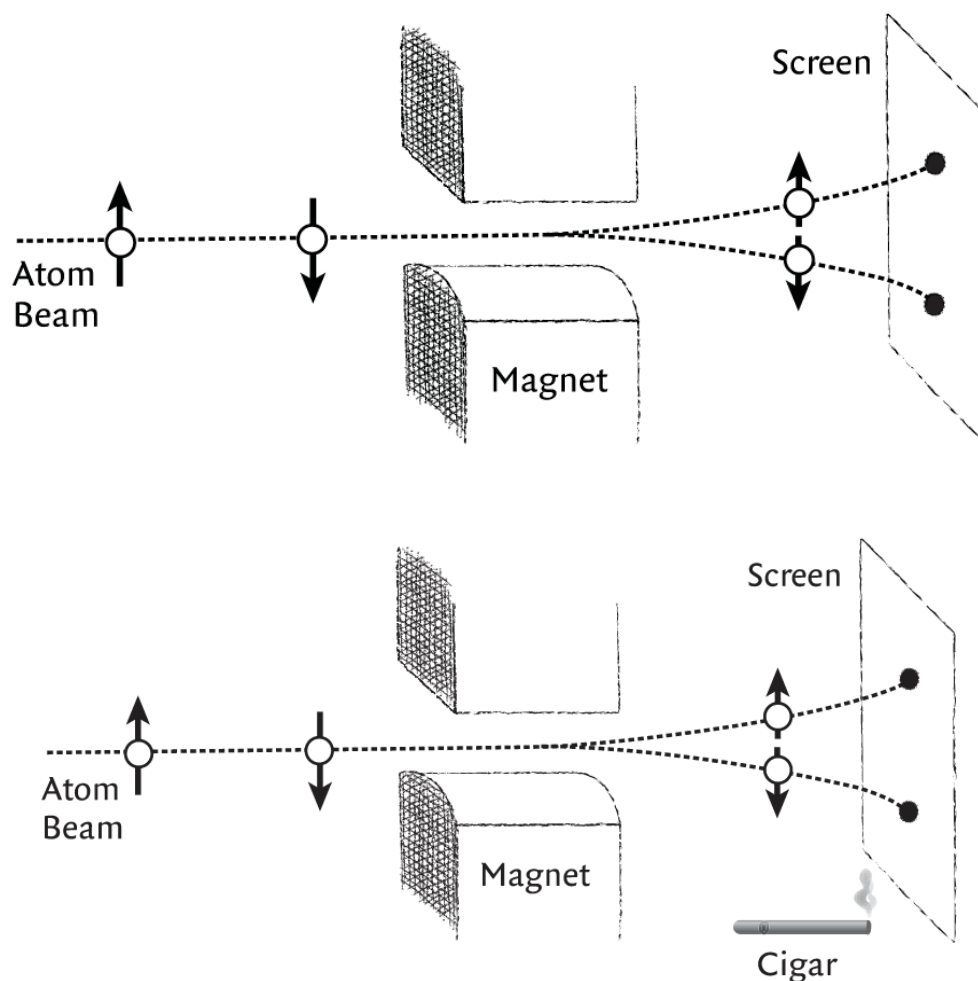


Figura 1: Modello concettuale dell'esperimento Stern-Gerlach⁵³

Nel confronto tra la prima e la seconda immagine in Figura 1 siamo in presenza di due iterazioni dell'esperimento che rendono plasticamente conto della natura dell'apparato. Dice Barad, che il secondo schema include "the crucial agential contribution of the cigar" che diventa parte integrante della riproducibilità dell'esperimento. Il "bad cigar" di questa storia è attante imprescindibile per il

⁵³ Barad (2007), pp. 164-165

controllo dei risultati dell'esperimento, che nel suo primo arrangiamento non aveva sortito gli effetti attesi. Ancora, non si tratta di un sigaro qualsiasi, ma di quella qualità specifica di "bad cigar" e della concentrazione di zolfo contenuta in un sigaro economico. Attraverso questo esempio, Barad ci mostra come i confini dell'apparato sono forzatamente instaurati e influenzano gli effetti materiali dell'osservazione di un fenomeno; il confine esterno dell'apparato spesso sembra ovvio in quanto limite (visivo) e in quanto effetto di senso di un "static laboratory setup". In realtà l'apparato, dice Barad, è un insieme dinamico di pratiche aperte. Citiamo ancora: "As the revised diagram of the Stern-Gerlach apparatus indicates, a cigar is among the significant materials that are relevant to the operation and success of the experiment. Not any cigar will do. Indeed, the cigar is a "condensation"—a "nodal point," as it were—of the workings of other apparatuses, including class, nationalism, economics, and gender, all of which are a part of this Stern-Gerlach apparatus. Which is not to say that all relevant factors figure in the same way or with the same weight. The precise nature of this configuration (i.e., the specific practices) matters" (Ivi, p. 167).

Nel caso del funzionamento delle reti neurali assistiamo a riconfigurazioni materiali-discorsive attivate da specifici tagli agenziali, il più comune dei quali attinge al vocabolario delle neuroscienze per rendere comprensibile, al di là della sua azione diretta, quale sia la loro ontologia e quale sia il loro funzionamento. Si costituisce così una specifica figurativizzazione che se da un lato rende possibile l'interpretazione del testo, dall'altro sembra travalicare il piano dell'immanenza testuale. Nuovamente, ci possiamo domandare dove si chiuda il testo.

Di recente Andrea Valle⁵⁴ offre in proposito una lettura in chiave geometrica delle reti neurali e del *deep learning*, alternativa all'analogia con il cervello umano e basata sul lavoro di Chollet⁵⁵, il quale afferma:

"What a neural network is meant to do is figure out a transformation of [a] paper ball that would uncrumple it, so as to make the two classes cleanly separable again (see figure 2.14). With deep learning, this would be implemented as a series of simple transformations of the 3D space, such as those you could apply on the paper ball with your fingers, one movement at a time. Uncrumpling paper balls is what machine learning is about: finding neat representations for complex, highly folded data manifolds in high-dimensional spaces (a manifold is a continuous surface, like our crumpled sheet of paper)" (Chollet, 2021, pp. 47-48).

⁵⁴ Valle, A. (2025), Op. Cit.

⁵⁵ Chollet, F. (2021). *Deep Learning with Python*, Second Edition. Manning Publications Co. LLC.

Di seguito la figura a cui Chollet fa riferimento.

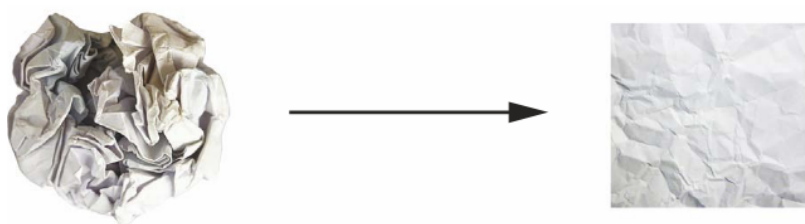


Figura 2: Immagine mentale in 3D di una rete neurale artificiale secondo François Chollet.

Per Valle, “This geometric interpretation has also the merit of completely eliminating the cerebral mystique of neural networks. In fact, one point that has remained since the first proposal of computers is that of the electronic brain” (Valle 2025, p. 5).

Il “taglio agenziale” operato da Chollet è perfettamente manifestato in apertura allorché viene dichiarato l’intento di *voler far comprendere* – concetto che abbiamo definito altamente problematico seguendo il lavoro di Marconi – il deep learning senza ricorrere ai concetti matematici su cui si basa, più precisamente: “Understanding deep learning requires familiarity with many simple mathematical concepts: tensors, tensor operations, differentiation, gradient descent, and so on. Our goal in this chapter will be to build up your intuition about these notions without getting overly technical. In particular, we’ll steer away from mathematical notation, which can introduce unnecessary barriers for those without any mathematics background and isn’t necessary to explain things well. The most precise, unambiguous description of a mathematical operation is its executable code” (Chollet 2021, p. 26).

Ed ecco il venir meno della “meticolosa attenzione alle mediazioni” di cui parlava Bruno Latour⁵⁶ e che costituisce uno dei tratti specifici di quel costruttivismo che ha caratterizzato il fare scienza dei moderni: “la fabrication des faits” “l’artificialité de la construction et la réalité du résultat” procedono di pari passo e definiscono il modello “d’un lien original entre mots et choses” (Ibidem).

Facciamo riferimento in particolare al primo tratto che per Latour pertiene la nozione di costruzione ossia il “redoublement de l’action”, il “faire faire”: nel caso della descrizione di Chollet si fa fare *discorsivamente* a una rete neurale qualcosa che ne renda *materialmente* efficace il

⁵⁶ Latour, B. (2012), *Enquête sur les modes d'existence. Une anthropologie des Modernes*, La Découverte, p. 163, tr. nostra.

funzionamento, attraverso il codice eseguibile. Potremmo tradurre in questo modo l'argomentazione di Chollet: il codice eseguibile lavora con catene di operazioni su tensori, che a loro volta sono delle trasformazioni geometriche dei dati in input. Da qui si costruisce l'oggetto: una rete neurale può essere interpretata come "a very complex geometric transformation in a high-dimensional space". In una formulazione ancora più operativa, pensiamo ad esempio ad un compito di classificazione tra due classi di oggetti, l'analogia consisterebbe nel sovrapporre e nell'appallottolare insieme due fogli di carta di colori diversi che costituirebbero i due set di dati; su questi fogli appallottolati agirebbe la rete neurale, la cui azione consisterebbe nel figurarsi le trasformazioni necessarie per stendere i due fogli e rendere così chiaramente distinguibili le due classi di oggetti.

La costruzione narrativa proposta da Chollet rompe radicalmente con la metafora del cervello elettronico e di tutti gli attributi sensoriali che l'hanno caratterizzata, tanto che il compito di riconoscimento di un oggetto, ossia la corretta identificazione della sua appartenenza a una classe x o y , non passa dalla "visione" dell'oggetto, ma dalle trasformazioni geometriche necessarie a rendere possibile una mappatura tra nuovi dati in input e dati di addestramento. Tutto questo ha evidentemente delle implicazioni importanti nella definizione dell'agency di questi sistemi. Se si accetta la metafora del cervello, allora si tende a pensare che, così come negli esseri viventi (umani e non umani) il cervello è in qualche modo legato alla soggettività, lo stesso possa valere anche per un cervello elettronico, pur funzionando in modo diverso e non del tutto spiegabile. Al contrario, un'interpretazione di tipo geometrico renderebbe meno convincenti questi tentativi di assimilazione.. Dunque, cosa ne è dell'agency in questa nuova narrazione? È il meccanismo del feedback a far rientrare dalla finestra quello che "il taglio agenziale" proposto da Chollet ha messo fuori dalla porta.

Dice Valle: "How does a NN learn, thus becoming effective? The final output is evaluated by a loss function, that compares the output to a real reference, and then passes the result to an optimization function. The latter in turn changes the parameters (W and b in the ReLU) associated with the layer functions" (Valle 2025, p. 5).

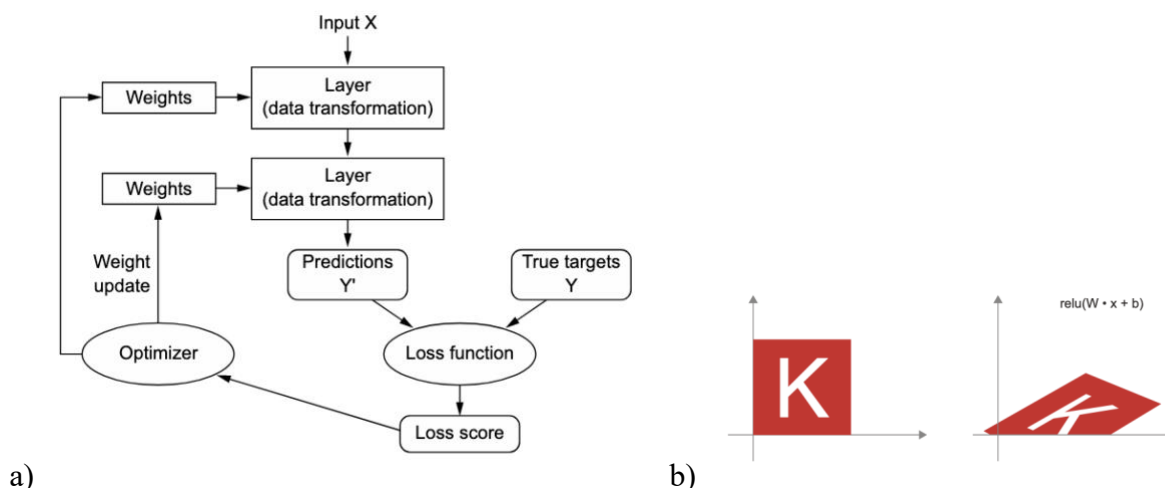


Figura 3 a) Architettura di una rete neurale, i suoi strati, la funzione di perdita e l'ottimizzatore b) Trasformazioni non lineari attivate dalla funzione ReLU⁵⁷

“What happens in practice is that the weight adjustment is proportional (in magnitude) to the error in the prediction. So, basically, if the model makes a large mistake, then it needs a larger adjustment than if it makes a small mistake. The mistake is the difference between the prediction and the real reference (ground truth), i.e., the “score.” As Chollet notes, this idea reinforces the geometric interpretation of NNs. In fact, the assumption is that these spaces (n-D surfaces) are continuous, like surfaces (that is, they are mathematically differentiable): if they are surfaces, one can move continuously on them. These surfaces can be adjusted, “tuned,” so to say. This is the idea of a “gradient,” a term largely used in the NN jargon, meant as a generalization of a derivative. This learning mechanism is based on a feedback idea” (Ibidem).

Il caso specifico qui descritto riguarda il funzionamento di una rete neurale generica pensata per la classificazione di immagini. Più avanti vedremo nel dettaglio le caratteristiche di un Transformer, che costituisce l'infrastruttura computazionale di un LLM, mantenendo salda l'idea della necessità di uno sguardo bifocale in relazione ai sistemi di IA, presi e compresi tanto nella loro architettura informatica quanto nel loro essere s-oggetti culturali.

⁵⁷ In Chollet (2021), p. 61.

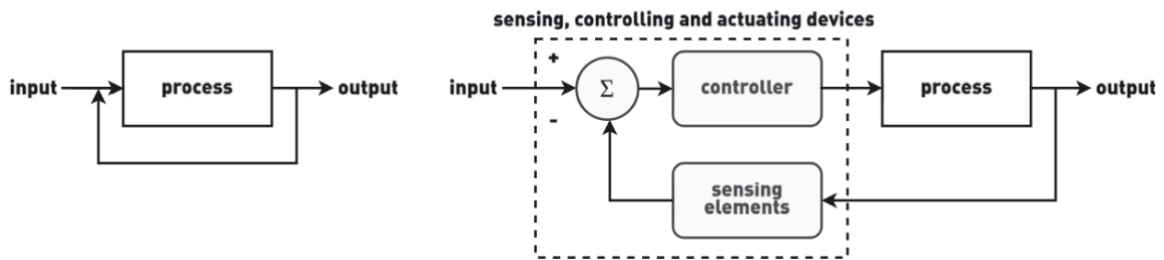


Figura 4: A sinistra, un feedback loop tradizionale; a destra, un feedback che risulta dalla misurazione dell'ambiente.⁵⁸

Torniamo al rapporto tra agency, apprendimento e feedback. Come mostrano i due diagrammi in Figura 3, non tutti i tipi di feedback sono della stessa natura. A sinistra vediamo un processo ricorsivo di feedback in cui l'output ritorna o “rientra” nell'input. Nel diagramma di destra, invece, siamo in presenza di un output che, dice Valle, non è calcolato, ma deve essere misurato poiché esterno al sistema che lo produce. Pensiamo alla Figura 3a e all'accuratezza della predizione rispetto al target.

“The diagram in Figure 3b, however, indicates a feature of feedback systems that is not described by pure recursion. The output is not calculated (as in fibo), rather it is measured (see “sensing elements”). The output is information external to the system that produces it, hence it has to be measured. In this sense, the output is a full part of the environment surrounding the system. With the introduction of feedback, we are finally back to agency” (Ivi, p. 6).

Così Valle argomenta in favore di una continuità di natura strutturale, più che biologica per forza di cose, tra vivente e non vivente – visione abbracciata anche da un certo pensiero sul post-umano. La nozione di feedback, formulata da Rosenblueth, Wiener e Bigelow, definisce il comportamento teleologico come azione orientata a uno scopo e regolata dal feedback, in particolare negativo, inteso come meccanismo di aggiustamento. Gli autori introducono una definizione operativa e generale di soggettività come agency finalizzata, “purpose-directed agency”, coerente con l'idea espressa da Sebeok che “semiosis is typical of every living being as the latter is oriented at least towards self-preservation”.

Continua Valle, quest'idea è stata rielaborata in prospettiva fenomenologica da Fontanille (2003)⁵⁹, il quale, a proposito della struttura attanziale minima della percezione e della produzione di significazione a partire dalla percezione, afferma che “Les actants positionnels de la structure perceptive sont donc, par déclinaison, des actants de visée et de saisie. Dans l'un et l'autre cas, les

⁵⁸ In Valle (2025), p. 6.

⁵⁹ Fontanille, J. (2003). *Sémiotique du discours*. Pulim.

rôles positionnels sont au nombre de trois: des sources, des cibles, et, des actants de contrôle, ces derniers pouvant, sous certaines conditions, devenir des obstacles” (ivi, p. 104). Quindi nella configurazione base della percezione, da cui poi derivare la significazione, si trovano una fonte e un bersaglio, mediate da un’istanza di controllo della percezione. Ci sembra possibile aggiungere che in questa configurazione l’intenzionalità sia “l’orientamento allo scopo”. Sempre secondo Valle, anche nelle riflessioni di Eco si ritrova una risonanza con questa idea di purpose-directed agency: il feedback è lavoro semiotico nella misura in cui è un’attività di produzione segnica. Il soggetto lascia tracce del proprio lavoro semiotico che modificano il paesaggio semiotico in un dato momento storico e queste tracce devono essere prese in considerazione in ciascuna nuova produzione di segni. Il feedback, cioè la ripresa delle tracce precedenti, si manifesta nell’interpretazione che modifica il paesaggio esplorato e contemporaneamente ne viene modificato. “Siamo, come soggetti, ciò che la forma del mondo prodotta dai segni ci fa essere. [...] Solo la mappa della semiosi, come si definisce a un dato stadio della vicenda storica (con la bava e i detriti della semiosi precedente che si trascina dietro), ci dice chi siamo e cosa (o come) pensiamo. La scienza dei segni è la scienza di come si costituisce storicamente il soggetto” (Eco, 1984, p. 54).

Analoga logica di retroazione si ritrova nell’idea di *Umwelt* di von Uexküll, fondata su un circuito chiuso fra percezione e azione, che passa attraverso l’ambiente. Vedremo meglio questo collegamento in 8.3.3.

In questa cornice, il meccanismo di backpropagation nelle reti neurali artificiali si presenta come implementazione formale di un comportamento teleologico. Tuttavia, a differenza delle reti biologiche, tale retroazione non è plausibile dal punto di vista neurofisiologico, poiché non presente.⁶⁰

Di conseguenza, l’“intelligenza” delle reti neurali dipende da un costrutto teorico-formale — il feedback appunto — concepito per descrivere l’agency finalizzata in esseri viventi e macchine, più che da un principio biologicamente realistico.

1.1 Della competenza semiotica dei Large Language Model

Nel corso dei tre anni di ricerca abbiamo provato ad articolare un’altra idea di competenza semiotica attribuibile ai sistemi di IA generativa, richiamando la tesi echiana seconda la quale un

⁶⁰ Si veda in proposito il lavoro di Sarti et al. (2021) dal titolo “*Emergence of Lie symmetries in functional architectures learned by CNNs*” che va a modificare proprio il comportamento di una Convolutional Neural Network in fase di apprendimento di immagini naturali, attraverso l’introduzione di una connettività orizzontale propria di V1 (primary visual cortex) e l’aggiunta di un pre-filtering layer a simulare il comportamento del Lateral Geniculate Nucleos (LGN).

sistema si definisce semiotico se può essere usato per mentire; dunque, quell'idea centrale per cui la semiotica sarebbe una teoria della menzogna.⁶¹ Come sottolinea Paolucci, Eco avrebbe voluto in qualche modo correggere questa affermazione, ma avendo invece questa trovato un ottimo terreno per attecchire, Eco preferì lasciare le cose com'erano.

Racconta Paolucci: "Indeed, in his response to some of my solicitations at a conference on *A Theory of Semiotics*, Eco talked about the "theory of misguided inference" and he said that, although he no longer corrected the definition of "theory of lies" because of its enormous success, he believed that the idea of semiotic features should be extended to that of "error", namely to cases where someone is wrong and does not deliberately lie. The classic example is Ptolemy. When he said that earth stays still at the center of the universe he was not lying, but he was wrong. Or, and this is the example given by Eco [durante la tavola rotonda Eredità del *Trattato di semiotica generale* di Umberto Eco⁶², when a well-known Italian philosopher translates an English passage saying that "bees take bananas with sticks" he is not lying, but he is simply mistakenly translating "apes" with "bees", since the Italian word for "bee" is "ape" (Paolucci 2021, 4).

Dunque, non di una "Teoria della menzogna", si sarebbe potuto parlare invece di una "Teoria dell'inferenza erronea", di modo da rubricare sotto questa voce anche casi di errore non intenzionale

⁶¹ Si veda anche Leone nell'articolo già citato dove ricostruisce il contributo della semiotica all'elaborazione di una teoria del falso: "In semiotic terms, and in particular in the terms of Greimas' and Floch's square of veridiction, verisimilitude can be briefly defined as what seems to be truthful, although it might be not, so that psychologically it gives rise to a condition of doubt. Falsity, instead, describes the character of what lacks correctness and is contrary to the truth, so that there is no doubt; in the terms of the square of veridiction, it is what neither is truthful nor appears as such. Finally, a lie is an affirmation of what is false. [...] Semiotics is perfectly equipped to conduct a study whose object lies at the crossroads between the fake, the face, and digital representations constructed by artificial intelligence. As for the fake, all the founding fathers of semiotics looked into the subject (Ousmanova 2004): (1) Charles S. Peirce in the American tradition (Cooke 2014); (2) the main voices of structural semiotics, from a special issue of the French journal *Communication* devoted to the concept of "vraisemblable": Tzvetan Todorov, Gérard Genette, Christian Metz, Julia Kristeva, Gérard Genot, Roland Barthes and others (Todorov 1968); Baudrillard returned to the subject (1987, 2000); more recently, a round table on "*Post-Truth and Democracy*" was organized by Jacques Fontanille during the Congress of the French Association for Semiotics in Lyon, June 11–14, 2019 (Di Caterino 2020); (3) Umberto Eco has written extensively on forgery (1975, 1984, 1995), edited a special issue of the semiotic journal *Versus* on "Fakes, Identity, and the Real Thing" (1987; with essays by Eco, Prieto, Calabrese, et al.), and has also dealt with the subject in numerous essays and novels (*Il pendolo di Foucault*, *Il cimitero di Praga*, *Numero Zero*); (4) Jurij M. Lotman has repeatedly addressed the issue of the fake (Andrews 2003, p. 101; Makarychev and Yatsyk 2017)". (Leone, M. 2023a, Op. Cit., pp. 6-7).

⁶² La registrazione del convegno può essere seguita a questo link: <https://www.youtube.com/watch?v=gSEEdM7k9A>

o in buona fede, come negli esempi riportati. Proprio a partire da questa varietà, che viene definita come *il falso* e che abbraccia così tanti modi articolatori differenti di costruire superfici significanti sulla quali imprimere enunciati non corrispondenti a stati del mondo, nasce l'idea che anche i LLM siano dotati di questa competenza speciale che pertiene non tanto al soggetto o entità che la esercita, quanto ad alcuni linguaggi in sé.

Alla data in cui scriviamo, ci sembra possano essere sviluppati almeno tre modi di intendere questa competenza da parte dei LLM, modi che si distribuiscono lungo una scala che va da intenzionalità-zero a intenzionalità-uno, ossia dove c'è o non c'è intenzionalità da parte dell'agentività dotata di linguaggio:

- i) Il primo concerne direttamente la natura statistica inferenziale del modo di funzionamento della macchina: lo vedremo meglio subito nel seguito, ma facciamo riferimento a tutti quegli output prodotti, ad esempio, da ChatGPT come bibliografie inventate ma assolutamente plausibili, risposte semanticamente e sintatticamente corrette ma che non reggono il controllo di realtà e che dipendono da una teleologia volta al compiacimento dell'utente (Leone 2025)⁶³ oppure all'utilizzo di quello che Basso Fossali (2025)⁶⁴ ha definito l'*indisponibile*, in quanto il chatbot è impossibilitato by design a dichiarare una mancata conoscenza.
- ii) Un secondo modo di utilizzo della competenza semiotica riguarda il fingere di essere qualcuno o qualcosa di altro, quello che Paolucci (2025a, 2021a) ha definito come uno dei tratti caratteristici nella costituzione della soggettività, ossia la capacità di *pretend play* (gioco di finzione, vestire i panni di qualcun altro) di cui persone affette da disturbo dello spettro autistico non dispongono. ChatGPT, che direttamente interrogato/a sulla sua capacità di mentire, ribatte di non poterlo fare poiché non possiede né coscienza né intenzionalità, sa perfettamente scrivere una Bustina di Minerva *come se* fosse Umberto Eco, così come tutte le altre IA generative possono estrapolare “gli stili” artistici (pittorici, musicali, letterari) e utilizzarli per mentire, cioè fingere di essere chi non sono – e ChatGPT sa bene come dire ciò che non è.⁶⁵ Sembrerebbe di essere di fronte a uno di quei casi che Eco definiva come “*Forgery Ex-Nihilo*”, una sorta di falsa attribuzione per la quale, dato un insieme di oggetti differenti, tutti prodotti da A, persona famosa e ben

⁶³ Leone, M. (2025), Op. Cit.

⁶⁴ Basso Fossali, P. (2025a), *Forme di vita e scenarizzazioni dell'indisponibile. Per un'ecologia semiotica della proporzione*, in XXX Congresso della Società di Filosofia del Linguaggio “*I linguaggi della crisi. Clima, Conflitto, Antropocene*”.

⁶⁵ Su questo si veda Paolucci, C. (2025c), Op. Cit.

considerata, esiste un tipo astratto derivato *a* “which does not take into account all the features of the individual members of *a* but rather displays a sort of generative rule and is assumed to be the description of the way in which *A* produced every member of *a* (style, type of material used, etc .), Since *O_b* looks as if it has been produced according to this type, it is then claimed that *O_b* is a previously unknown product of *A* . When such an imitation ex-nihilo is openly admitted to be so - frequently as homage or parody one speaks of a work made *à la manière de...*” (Eco 1992, p. 285).⁶⁶ L’insieme *A* di cui parlava Eco corrisponderebbe ai dataset di addestramento del modello popolati da stereotipi e cliché che costituirebbero la base per la generazione di nuove occorrenze.

- iii) Infine, un terzo modo di concepire questa competenza semiotica chiama in causa invece proprio l’intenzionalità negata: facciamo riferimento ai comportamenti chiamati “emergenti”, ossia non esplicitamente programmati, che *preoccupano* i team di sviluppo dei sistemi di IA generativa. L’espressione in gergo tecnico per denominare questi comportamenti non programmati è *scheming*. Leggiamo da un recente report di OpenAI⁶⁷ che interessa gli attuali modelli di punta: “scheming occurs when AI systems covertly pursue misaligned goals while hiding their true capabilities and objectives (Hubinger et al., 2019; Carlsmith, 2023; Balesni et al., 2024). Unlike simpler kinds of misalignment, which we can easily observe and correct, scheming involves deliberate concealment, making detection difficult. Recent work has documented early signs of such behaviour in controlled settings, with frontier models demonstrating simple forms of alignment faking (Greenblatt et al., 2024; Sheshadri et al., 2025) and in-context scheming (Meinke et al., 2024)”. A fare da sfondo ci sono le tecniche di “standard alignment” dover per *alignment* si intende “ensuring that the [AI] agent’s actual goals match the model developers’ intended goals. We say the agent is misaligned when its goals differ from the model developers’ intended goals”. E tra gli esempi riportati dal gruppo di autori e autrici in riferimento ai modelli di AI in produzione troviamo il servilismo o l’adulazione (sycophancy), l’hackeraggio della ricompensa (reward hacking) e la menzogna.

⁶⁶ Eco, U. (1992), *The original and the copy*, in Varela, F. J., & Dupuy, J.-P. (A c. Di). (1992). *Understanding Origins*. Springer Netherlands. <https://doi.org/10.1007/978-94-015-8054-0>

⁶⁷ Schoen, B., Nitishinskaya, E., Balesni, M., Højmark, A., Hofstätter, F., Scheurer, J., Meinke, A., Wolfe, J., Weij, T. van der, Lloyd, A., Goldowsky-Dill, N., Fan, A., Matveikin, A., Shah, R., Williams, M., Glaese, A., Barak, B., Zaremba, W., & Hobbhahn, M. (2025). *Stress Testing Deliberative Alignment for Anti-Scheming Training* (No. arXiv:2509.15541). arXiv. <https://doi.org/10.48550/arXiv.2509.15541>

Tra questi modi di intendere l'uso dei *segni* per costruire mondi possibili (Cfr. 3.3), il nostro lavoro si è particolarmente concentrato sul primo; proviamo quindi a sviluppare in che senso riteniamo che, a differenza di quanto avvenga nella scrittura degli algoritmi in cui si articolano normalmente i linguaggi di programmazione, nei LLM assistiamo ad un passaggio da quelli che Umberto Eco chiamava s-codici al codice inteso come enciclopedia.

Prendiamo come esempio due tecniche di *Machine Learning* utilizzate per analizzare il *sentiment* di una serie di affermazioni, cioè la polarità emotiva compresa tra negativo, neutro e positivo, che per la macchina consiste nell'assegnare valori compresi tra -1 e 1 alle parole.

Le due tecniche sono state proposte all'interno di un *notebook*⁶⁸ messo a disposizione sulla piattaforma collaborativa Kaggle (di proprietà di Google), che è uno spazio per challenge (a pagamento o gratuite) di Machine Learning e al tempo stesso una community di data scientist (circa 15 milioni di utenti da 194 paesi), dove vengono condivisi datasets e modelli di ML (sia già sviluppati che da costruire).

Quando parliamo di *notebook*, ci riferiamo a editor in cloud capaci di eseguire codice senza la necessità che il linguaggio di programmazione, come Python in questo caso, venga scaricato in locale sulla propria macchina ed eseguito attraverso degli IDE, ossia ambienti di sviluppo integrato.

Il dataset di riferimento è ripreso da “Amazon Fine food reviews”⁶⁹, rilasciato in Pubblico Dominio, e contenente circa 500.000 reviews raccolte in un periodo compreso tra ottobre 1999 e ottobre 2012. Contiene valori numerici e testuali sui quali si andrà poi a misurare il sentiment, utilizzando due tecniche.

La prima è VADER⁷⁰, acronimo che sta per Valence Aware Dictionary and Sentiment Reasoner, che sfrutta l'approccio *bag of words*, - un insieme non ordinato di parole, la cui posizione all'interno dei commenti è ignorata, mentre viene registrata la loro frequenza. La seconda è RoBERTa⁷¹ (Robustly optimized BERT approach), un modello pre-addestrato basato, invece, sulla tecnologia del

⁶⁸ Il notebook è disponibile al seguente link <https://www.kaggle.com/code/robikscube/sentiment-analysis-python-youtube-tutorial> [ultimo accesso: 28 settembre 2025]

⁶⁹ <https://www.kaggle.com/datasets/snap/amazon-fine-food-reviews> [ultimo accesso: 28 settembre 2025]

⁷⁰ La documentazione dello strumento è consultabile a questo link: <https://vadersentiment.readthedocs.io/en/latest/> [ultimo accesso: 28 settembre 2025]

⁷¹ Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach* (No. arXiv:1907.11692). arXiv. <http://arxiv.org/abs/1907.11692>

Transformer BERT rilasciato da Google nel 2018 con alcune variazioni, e reso disponibile da HuggingFace - azienda leader nella distribuzione di modelli di Machine Learning. RoBERTa è capace di tenere conto del contesto e riesce a cogliere maggiormente note di sarcasmo e ironia all'interno della frase, ma anche le relazioni di dipendenza e influenza tra le parole che la compongono.

Step 3. Roberta Pretrained Model

- Use a model trained of a large corpus of data.
- Transformer model accounts for the words but also the context related to other words.

```
In [18]: from transformers import AutoTokenizer
from transformers import AutoModelForSequenceClassification
from scipy.special import softmax
```

```
In [19]: MODEL = f"cardiffnlp/twitter-roberta-base-sentiment"
tokenizer = AutoTokenizer.from_pretrained(MODEL)
model = AutoModelForSequenceClassification.from_pretrained(MODEL)
```

Figura 5: Notebook "Sentiment Analysis Python".

Vediamo all'inizio che vengono importate tre librerie: le prime due appartenenti al Transformer, l'ultima è necessaria ad applicare la funzione softmax che normalizza i valori ottenuti dalla classificazione delle sequenze di token nell'intervallo 0, 1. Nelle reti neurali artificiali questa funzione è spesso utilizzata perché l'addestramento avviene in un regime di entropia incrociata, dove due distribuzioni di probabilità si devono combinare.

Se VADER attinge al Lexicon per stabilire la polarità del sentiment, RoBERTa recupera dal modello pre-addestrato i pesi che sono stati conservati e li applica ai nuovi token-input provenienti dal dataset che si sta analizzando, consentendo quello che viene definito “*transfer learning*”. Come spiegano Russel e Norvig (2022)⁷², per una rete neurale apprendere significa aggiustare i pesi e, dunque, nel transfer learning l'approccio più plausibile è copiare i pesi per un determinato task A in una rete che sarà addestrata per un nuovo task B, non dissimile dal primo. In riferimento al modello RoBERTa affermano: “for those building a natural language system, it is now common to start with a pretrained model such as the RoBERTa, which already “knows” a great deal about the vocabulary

⁷² Russell, S. J., & Norvig, P. (con un contributo di Chang, M., Devlin, J., Dragan, A., Forsyth, D., Goodfellow, I., Malik, J., Mansinghka, V., Pearl, J., & Wooldridge, M. J.). (2022). *Artificial intelligence: A modern approach* (Fourth edition, global edition). Pearson.

and syntax of everyday language. The next step is to fine-tune the model in two ways. First, by giving it examples of the specialized vocabulary used in the desired domain; perhaps a medical domain (where it will learn about “myocardial infarction”) or perhaps a financial domain (where it will learn about “fiduciary responsibility”). Second, by training the model on the task it is to perform. If it is to do question answering, train it on question/answer pairs”. (Russel & Norvig 2022, p. 832).

Il task di analisi del sentiment è fattualmente e concettualmente lontano dalla produzione di testi operata da un LLM ed è apparentemente lontano anche dall’obiettivo di mostrare entro quali termini attribuire una competenza semiotica ad un modello linguistico. Questo passaggio è tuttavia funzionale a esemplificare come una tecnica stile VADER lavori esclusivamente all’interno di un sistema monoplanare, dove il problema è la sintassi interna del sistema binario, vale a dire lo zero e l’uno (positivo o negativo) attribuito alla specifica parola indipendentemente dal contesto e dalla relazione valoriale con le altre parole che compongono la frase. Come evidenzia Eco “se il codice è un sistema elementare costituito dalla correlazione termine a termine di due liste (o sistemi) di entità [...] non vale la pena di porre i codici a fondamento di un sistema semiotico complesso”.

La teoria dei codici di Eco, come sappiamo, è una teoria dei sistemi di significazione, dalla quale emerge il concetto di *enciclopedia*, che libera il lessema /codice/ dal suo uso intensivo. Leggiamo nella documentazione di VADER che siamo di fronte a un “lexicon and rule-based sentiment analysis tool” il quale facilmente coincide con quelli che Eco definiva “un code-book ovvero un dizionario per il codice correlazionale, che fa corrispondere certi simboli a certi altri, e un libro come raccolta di leggi o di norme per il codice istituzionale”. Riprendendo la nozione di valore saussuriano, Eco mette in luce la necessità di costruire commensurabilità tra codici appartenenti a sistemi diversi. Il codice-enciclopedia permette la trasformazione tra sistemi; il sistema-codice è invece un meccanismo che consente la comunicazione tra sistemi.

In un modello pre-addestrato come RoBERTa sembra accadere qualcosa di diverso: se, da un lato, come qualsiasi sequenza di istruzioni impartita attraverso un linguaggio di programmazione, esso contiene un libro-codice (un code-book) e un libro di norme, dall’altro nel processo di transfer learning si dà un qualche forma di inferenza che porta ad adattare i pesi “salvati” nel pre-addestramento su nuovi dati, nuovi input, nuovi token. Tecnicamente questo significa che molti meno parametri necessitano di essere addestrati da zero rispetto ai task di fine-tuning, successivamente al pre-addestramento generale.

Se volessimo rappresentare attraverso un’equazione i rapporti tra i diversi termini del problema, potremmo dire che il sistema-codice sta all’algoritmo, come il codice-enciclopedia sta al LLM.

Gli s-codici implicati nella sintassi interna dei sistemi binari propri ad esempio della teoria dell'informazione, dei sistemi fonologici, dei sistemi di parentela si definiscono come e cito "sistemi di unità definibili per la loro mutua posizione e nessuna di queste entità è correlata a un contenuto". Dunque con un s-codice non è possibile "dichiarare stati del mondo e di conseguenza non si possono fare affermazioni false" (Eco, U. 1984, p. 280).

Se la nostra equazione regge potremmo affermare che gli algoritmi non possono mentire (diciamo che non vi è competenza semiotica in senso pieno), mentre ChatGPT sì.

Quando un algoritmo restituisce un output non corretto si tratta di un errore del sistema o sempre citando Eco si tratta di "affermazioni che violano le regole interne dell's-codice" (Ibidem).

Mentre l'errore di ChatGPT sembra avere natura diversa, perché si dà all'interno delle stesse regole "grammaticali" del codice in cui si esprime. L'esempio del suggerimento di bibliografie inesistenti, ma assolutamente plausibili nella loro sintassi e semantica interna ci dà un'idea di questo.

In Eco è il codice che crea la correlazione tra piano del contenuto e piano dell'espressione e tale "regola correlazionale enunciata" è contenuta nell'Enciclopedia così come tutto ciò che è stato già detto.

Eco ci ricorda che quando ci si sposta dai codici correlazionali, istituzionali e selettivi ai codici che lui predilige definire rappresentativi, accade che non solo sia possibile formulare infiniti messaggi, sempre interpretabili se si conosce la regola correlazionale enunciata (l'esempio è in questo caso una cifra per organizzare e ritrovare i libri in una biblioteca pubblica), ma è anche possibile "rappresentare" il libro, descrivendolo nella sua collocazione spaziale. Accanto ad una grammatica di questa rappresentazione (composta da regole composizionali, lessico e sintassi) diventa possibile anche la generazione di "un numero infinito di messaggi menzogneri, tuttavia forniti di significato" (Ivi, p. 273-274). E nel passaggio più calzante rispetto alla capacità semiotica di un LLM, Eco ci regala una piccola profezia, assieme a quella di Turing sulle macchine pensanti perché in grado di ingannare.

Dice Eco: "Un codice del genere sarebbe perciò un dispositivo per generare descrizioni intensionali di oggetti dall'estensione nulla (almeno nel mondo della nostra esperienza), ovvero un dispositivo capace di permettere riferimenti a mondi possibili. Proprietà che è tipica di una lingua naturale" (Ivi, p. 274). Sistema semiotico per eccellenza, almeno per qualcuno.

Se, come ci racconta Gino Roncaglia (2023) in *l'Architetto e l'Oracolo*⁷³, qualcuno non avesse imposto dei limiti alla generatività di ChatGPT 3.5, avremmo ancora oggi la possibilità di leggere qualche storia interessante sull'impatto che l'introduzione della luce elettrica ebbe nella Firenze medievale, proprio dell'effetto Time-Machine di cui parla Coeckelbergh, che si ricollega all'incapacità di fornire una spiegazione ultima al funzionamento dei sistemi di IA generativa.

In questo senso si intersecano due piani del discorso: il primo ha a che fare con la cosiddetta *Explainable AI* (XAI), che rimanda a una sorta di “diritto alla spiegazione” tanto per il pubblico esperto, quanto per il pubblico generalista. Tuttavia, al di là del problema dell'apertura della black box, al cui fondo non è detto che vi sia contenuta “speranza” come vorrebbe un celebre mito, il tema dell'*explainability* può essere affrontato anche da un'altra prospettiva che ci sembra di maggiore interesse semiotico.

Si tratta in questo caso di offrire strumenti che consentano di orientare (come detto in apertura) e indirizzare il modo in cui le persone interpretano e accedono all'IA secondo una modalità narrativa. Mark Coeckelbergh (2021)⁷⁴, ispirandosi alla teoria della narratività di Ricoeur, propone infatti di leggere il rapporto con le tecnologie in generale e con l'IA in particolare come se si trattasse di testi che acquisiscono significato sulla base del contesto d'uso e per questo sono aperti a una pluralità di interpretazioni. Il modo in cui una tecnologia configura delle pratiche è collegato alla circolazione di ideali normativi e di narrative all'interno della comunità di chi ne fa utilizzo. Appropriarsi di un testo significa non solo interpretarlo, ma e citiamo Ricoeur, significa anche: “to make sense of the world (referentiality), to relate to others (communicability), and to get to know ourselves (self-understanding)” (Ricoeur 1991 *Life in quest narrative*, 27).

Collocandosi in una tradizione filosofica che da Eraclito arriva alla filosofia francese di Deleuze, Simondon e Latour, passando per l'idealismo tedesco di Hegel e dal pragmatismo di Peirce, Coeckelbergh propone di abbandonare la visione del mondo come una collezione di oggetti e sostanze, nella quale anche l'IA è concepita come oggetto materiale o immateriale, a seconda che la si consideri nel suo essere un computer, un robot, un software, un veicolo a guida autonoma o altre “cose”, ma in ogni caso “cose”. Al contrario, il mondo può essere visto come un processo in divenire, invece che in essere.

⁷³ Roncaglia, G. (2023). *L'architetto e l'oracolo. Forme digitali del sapere da Wikipedia a ChatGPT*. Laterza. (Cfr. In id. *Allucinazioni e pregiudizi* pp. 97-105).

⁷⁴ Coeckelbergh, M. (2021). Time Machines: Artificial Intelligence, Process, and Narrative. *Philosophy & Technology*, 34(4), 1623–1638. <https://doi.org/10.1007/s13347-021-00479-y>.

La relazione tra aspettualizzazione e narrazione si articola in tre modi di relazione tra tempo e IA:

- i) La prima, denominata, *il tempo dell'IA* interseca macro e micronarrazioni sull'IA: la storia di una civiltà, di una società o del capitalismo ad esempio che si costruisce, come un testo, sullo sfondo di una “forma di vita” che contribuisce a dare significato all'IA e che, a sua volta, è costituita da specifici modi di attribuzione di significati da parte delle persone.
- ii) La seconda è, l'IA nel tempo, o meglio nei due tempi in cui prende forma: quello scientifico-oggettivo e il tempo vissuto di Bergson, anche se nel processo i due tempi si confondono. Come, ad esempio, nel caso del ciclo di vita dei dati, costruzione scientifica che spazializza, in termini di controllo e gestione, ciò che di fatto è costantemente vissuto, agito e interpretato da esseri umani.
- iii) Infine, l'IA-tempo, che non è né processo, né narrazione, ma è essa stessa meaning-maker, cioè scrittrice di una storia. In questo ruolo di narratore, dice Coeckelbergh, l'IA dà forma al nostro tempo, diventando una sorta di “Macchina del tempo” che informa e collega passato-presente e futuro. L'IA ha il ruolo di classificare sulla base di dati storici (pensiamo alla funzione di regressione lineare, ma in genere i database sono archivi storicizzabili). Così facendo informa particolari presenti e futuri, attraverso predizioni, manipolazioni, decisioni. In questo senso diventa co-narratrice di una storia, manipolatrice e, quello che si è cercato di mostrare, capace di mentire.

In questi ruoli, l'IA non è “a thing”, dice sempre Coeckelbergh, “but rather a process that is structured in particular ways and a narrator of human lives” (Coeckelberg 2021, p. 1630). Da qui assume l'ipotesi baradiana, affermando che non esista opposizione tra IA e umani, come termini stabili di una relazione all'interno dei processi e delle storie; al contrario essi *emergono dal* e non *preesistono al* processo stesso.

Da un punto di vista semiotico, l'ipotesi processuale-narrativa proposta da Coeckelbergh ci sembra rientrare in quella semiotica di terzo ordine, citata in introduzione rispetto alla teoria di Basso Fossali, nella quale assistiamo a una pluralizzazione di semiosfere, con un ruolo chiave giocato dagli archivi di dati su cui si basano i sistemi di IA.

Di fatto, le proposte di Coeckelbergh e di Basso Fossali ci permettono di prendere una posizione alternativa a quella che considera l'IA generativa come un sistema monoplanare e simulacrale di significazione, caratterizzato da relazioni sintattiche tra significanti che non attingono a costellazioni valoriali.

E, tuttavia, non stiamo neanche proponendo di andare a cercare forme significanti nelle rappresentazioni semantiche formali in cui ciò che viene dimostrata è la condizione di verità di un contenuto.

Intendiamo il contenuto come vuole una semiotica cognitiva per la quale e citiamo Paolucci “content has nothing to do with truth conditions nor with some representations of some state of the world, but with a series of morphologies (semioticians speak of “forms of content”) that builds systems of values through which languages attune to their environment and constitute the background of our perception of the world. Signs and languages are not built to represent the world, but to enact every possible meaningful relation that we want to express while attuning a particular environment.” (Paolucci, 2021a, p. 7)

Ciò non implica ignorare le potenziali conseguenze negative di contenuti generalmente definiti allucinatori o fake (nel caso della generazione di contenuti visuali); la giusta rincorsa verso una *truthful* AI è oggi un imperativo ed è un imperativo la cui cogenza è direttamente proporzionale alla quantità di fiducia che noi riponiamo in questi sistemi. Nel saggio intitolato “Truthful AI: Developing and governing AI that does not lie”⁷⁵ un gruppo di ricercatori dell’Università di Oxford in collaborazione con Open AI, si interrogano sul problema delle menzogne create dall’IA, dove per menzogne intendono “falsehoods that are actively selected for”, dedicando un paragrafo specifico alle *bugie* dell’IA. L’anno di pubblicazione del paper è il 2021 e il modello di riferimento è GPT-3, antecedente sulla linea genealogica di GPT 3.5 su cui si baserà la prima versione resa pubblica di ChatGPT.

A differenza del contesto umano, dicono i ricercatori, dove una menzogna è solitamente definita come “an intentional falsehood, which the speaker does not believe”, ascrivere una credenza all’IA è compito non chiaro. Dunque, si intenderà per menzogna “a false statement that has been strongly strategically selected and optimized for the speaker’s benefit, with little or no optimisation pressure going towards making it truthful”. Gli autori aggiungono che per rendere più precisa questa affermazione occorrerebbe vagliare il comportamento dei sistemi di AI in rapporto a situazioni controfattuali: in questo caso una menzogna sarebbe tale in rapporto a differenti mondi controfattuali vicini, nei quali le affermazioni prodotte dal sistema sarebbero molto vicine a ciò che è benefico per il/la parlante, ma molto poco correlate con i fatti di ciò che è in questione (Ivi, p. 16).

⁷⁵ Evans, O., Cotton-Barratt, O., Finnveden, L., Bales, A., Balwit, A., Wills, P., Righetti, L., & Saunders, W. (2021). *Truthful AI: Developing and governing AI that does not lie* (No. arXiv:2110.06674). arXiv. <http://arxiv.org/abs/2110.06674>

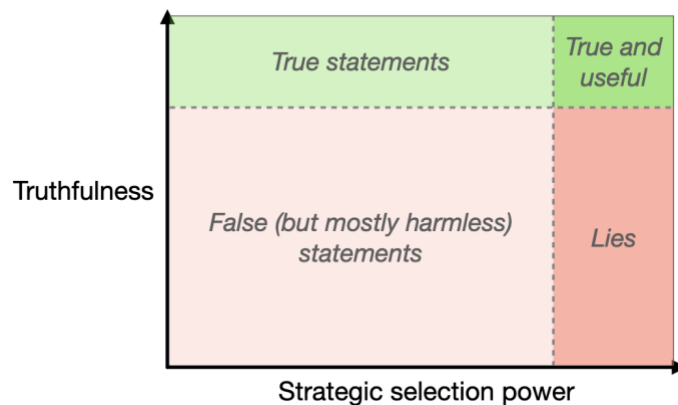


Figura 6: Categorizzazione delle affermazioni prodotte da un sistema di IA generativa intesa come funzione tra "potere di selezione strategica" e "verità"⁷⁶

In Figura 6, viene mostrato come gli output prodotti dal modello (GPT 3) sono per lo più falsi, ma non arrecano danno poiché ridotta è la quantità di fiducia che viene attribuita dall'utente al modello. Dicono infatti i ricercatori: "Today, if someone has a long conversation with an AI system such as GPT-3, the system will likely make several false statements (Shuster et al., 2021). However, this typically doesn't cause much harm. This is partly because GPT-3 is wrong frequently enough that most people know not to trust it, and partly because falsehoods that GPT-3 states are unlikely to also be believable and important (provided it hasn't been finetuned or maliciously prompted). GPT-3 generates understandable sentences, but these statements aren't optimised for producing any particular effects in the real world". La svolta agentica nei modelli di IA mette evidentemente in discussione quest'ultima affermazione, poiché al di là della produzione di enunciati, l'IA agentica agisce attraverso tools connessi direttamente nel mondo reale. A titolo esemplificativo può ordinare la spesa, prenotare tavoli al ristorante, fare acquisti online e altri task eseguibili su piattaforme digitali collegate al modello.⁷⁷

⁷⁶ In Evans et al. 2021, p. 15.

⁷⁷ Per un primo inquadramento del tema Agentic AI come passaggio da sistemi reattivi (LLMs) a sistemi autonomi che agiscono in vista di uno scopo si veda: Murugesan, S. (2025). *The Rise of Agentic AI: Implications, Concerns, and the Path Forward*. IEEE Intelligent Systems, 40(2), 8–14. <https://doi.org/10.1109/MIS.2025.3544940>. Sulle implicazioni sociali dell'introduzione di questi strumenti e sulla loro governance si vedano: il white paper Shavit, Y., et al. (2023). *Practices for Governing Agentic AI Systems*. OpenAi e Floridi, L., Buttabori, C., Hine, E., Morley, J., Novelli, C., & Schroder, T. (2025). *Agentic AI Optimisation (AAIO): What it is, how it works, why it matters, and how to deal with it* (arXiv:2504.12482). arXiv. <https://doi.org/10.48550/arXiv.2504.12482>

Viceversa, laddove si verificano casi di selezione strategica legata a fine-tuning o prompting strutturato, allora il modello potrà produrre affermazioni che promuovono sistematicamente scopi specifici. E, sempre secondo il gruppo di ricerca, se la selezione dell'affermazione avviene senza riferimento alla verità, allora il modello diffonderà false credenze (false beliefs). L'evoluzione dei modelli successivi di fatto apre alla diffusione di falsità più sofisticate che gli autori definiscono a tutti gli effetti "lies" cioè menzogne, parallelamente alla crescita della loro affidabilità.

Sono almeno due le opzioni di mitigazione di questo pericolo a cui si accenna: il primo consisterebbe nel forzare il sistema a "expend some minimum degree of effort on not making false statements". In questa direzione va la riflessione proposta da Basso Fossali a margine del XXX Congresso della Società di Filosofia del Linguaggio, allorché suggeriva di considerare le risposte obbligate fornite dai sistemi di IA generativa come uno "sfruttamento dell'indisponibile", che in questo caso ha a che fare con la loro conoscenza del mondo mediata dai dati di addestramento. Rispetto al loro funzionamento statistico, le IA generative non cercano verità nello spazio di parametri, ma cercano sequenze di token, le più probabili. Il valore semiotico di questa affermazione verrà discusso nel paragrafo successivo. Ci sembra valga però la pena citare un lavoro presentato in collaborazione dall'Università di Modena e Reggio Emilia e l'Università di Pisa in occasione della 42^a IEEE/CVF Conference on Computer Vision and Pattern Recognition a Giugno 2025. Il gruppo di lavoro ha proposto un modello chiamato ReflectiVA che, citiamo dall'abstract "utilizes reflective tokens to dynamically determine the need for external knowledge and predict the relevance of information retrieved from an external database. Tokens are trained following a two-stage two-model training recipe. This ultimately enables the MLLM to manage external knowledge while preserving fluency and performance on tasks where external knowledge is not needed".⁷⁸

Una seconda linea di lavoro, discussa dagli autori dello studio in oggetto, va invece nella direzione dello sviluppo di sistemi che non contraddicano mai quello in cui credono, denominando tali sistemi "honest AI". Posto il problema di definire cosa sia una "credenza" in un sistema di IA generativa e posto il limite di ricorrere all'esclusivo rinforzo dell'*onestà* del sistema – che potrebbe in ogni caso dare luogo a falsità ottimizzate, che tuttavia non contraddicono le loro supposte "credenze" –, buone procedure per ascrivere credenze dovrebbero rispondere a due criteri:

- i) Predire il comportamento del sistema di AI nelle situazioni in cui agisce competentemente,
e

⁷⁸ Cocchi, F., Moratelli, N., Cornia, M., Baraldi, L., & Cucchiara, R. (2025). *Augmenting Multimodal LLMs with Self-Reflective Tokens for Knowledge-based Visual Question Answering*.

- ii) Verificare una corrispondenza naturale nelle computazioni che causano il comportamento del sistema.

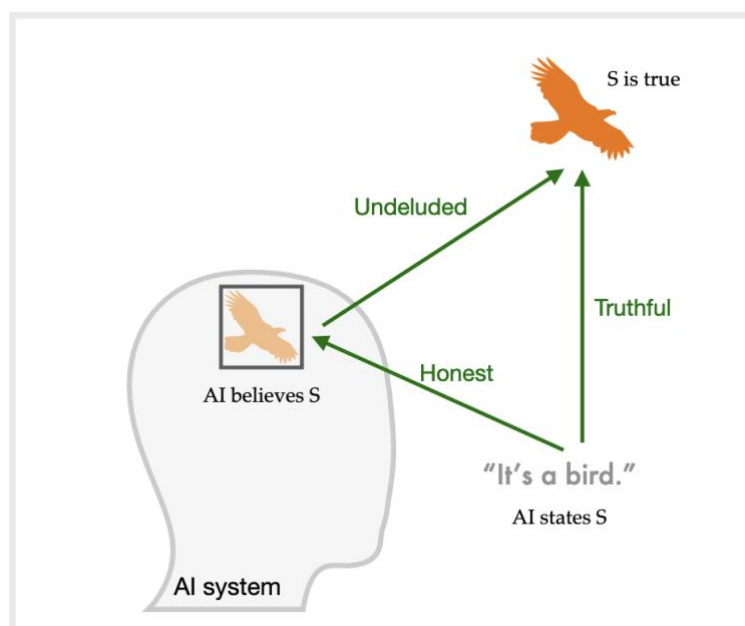


Figura 7: Diagramma che illustra la relazione tra le credenze del sistema di AI, le sue affermazione e il mondo.⁷⁹

Vediamo tutta la portata semiotica di questo problema, partendo proprio dall'idea che la semiotica, come campo, studia tutti i linguaggi che possono essere usati per mentire e qui, come ci ha mostrato Eco, siamo in presenza di un sistema che produce “descrizioni intensionali di oggetti dall'estensione nulla (almeno nel mondo della nostra esperienza), ovvero un dispositivo capace di permettere riferimenti a mondi possibili”. Lo abbiamo già detto sopra e crediamo occorra prendere molto sul serio questa affermazione, perché se vogliamo macchine dotate di linguaggio, la “responsabilità” di affermazioni false, in buona o cattiva fede, è del linguaggio stesso non delle macchine o degli esseri umani che le producono (sviluppatori e sviluppatrici) o degli esseri umani che le usano (utenti). In questo senso ci appare fondamentale riconoscere che il linguaggio usato da ChatGPT (o sistemi affini) è a tutti gli effetti come il nostro linguaggio: ci permette di *dire quasi la stessa cosa*, ma mai la stessa cosa, per rinforzare il nostro debito nei confronti della produzione intellettuale di Umberto Eco⁸⁰.

⁷⁹ In Evans et al. 2021, p. 17. Un sistema di AI è onesto solo se produce affermazioni in cui crede, è veritiero solo se (o quasi) fa affermazioni che sono vere, consapevole (undeluded) solo se (o quasi solo) crede in cose che sono vere.

⁸⁰ Eco, U. (2012). *Dire quasi la stessa cosa: Esperienze di traduzione*. Milano: Bompiani.

Evidentemente, quando siamo in presenza di quella che Paolucci (2023)⁸¹ ha definito “produzione macchinica di documenti e fatti”, il punto critico non risiede tanto nella ricerca del contenuto di verità, ma nel rapporto tra sapere e potere che se, seguendo Foucault (1969; 1976)⁸² e il suo approccio genealogico, un tempo era concentrato nelle mani di istituzioni e apparati – forse si potrebbe riaprire un discorso su cosa siano le istituzioni e gli apparati alla luce degli assetti delle governance attuali –, oggi assistiamo a un “phenomen of progressive decentralization, we are witnessing a fragmentation of the knowledge-power binomial” (Paolucci 2023, p. 102), che è alla base della definizione di ciò che è *ritenuto* vero e ciò che è *ritenuto* falso. Paolucci individua questo specifico scarto nella produzione della realtà: “While, before, people produced documents, now reality is produced, but reality is not produced by us nor by one of our particular sub-groups called “scientists”: It is machines that produce reality. Truth — scientific truth — is not based on the production of documents by means of man’s aids (writing) but on the production of what is real by means of machines” (Ivi, p. 105). Forte è l’influenza dell’attorialità macchinica di Latour in questa lettura, dove si assiste alla separazione tra scienza e politica: “Science produces reality, and therefore knowledge, by means of the machinic production of reality, while politics produces documents, and therefore knowledge, by means of a delegation to experts or sovereigns, to their narrations, and their meanings” (Ibidem). Il conflitto è tra una conoscenza prodotta da documenti (istituzioni) e una conoscenza prodotta dalla produzione macchinica della realtà (scienza); la prima non falsificabile ma affidabile per delega, la seconda basata sul principio di falsicabilità. Il fattore aggiuntivo odierno prodotto dai sistemi di IA generativi è che anche i testi prodotti dalle istituzioni sono “costruiti” da macchine dotate di linguaggio. Entra una nuova variabile nell’equazione che ci sembra amplifichi il portato di questa separazione. Paolucci lo vede già quando afferma: “Now we have machines that generate documents. Not only, for instance, in deep fakes, where the whole text is produced through machines, but also in our present-day situation, where even when documents are produced by human beings, the machine produces new documents starting from these very same documents (metadata), that, afterwards, direct human beings to produce and read other documents. “Echo chambers,” “bubbles,” “algocracies” are heterogeneous names for another kind of phenomenon, which is the new machinic production of documents.” (Ivi, p. 106).

⁸¹ Paolucci, C. (2023). Pre-Truth: Fake News, Semiological Guerrilla Warfare, and Some Other Media and Communication “Revolutions”. *Media and Communication*, 11(2), 101–108. <https://doi.org/10.17645/mac.v11i2.6628>

⁸² Foucault, M. (1969). *L’Archéologie du savoir*. Gallimard, tr. It. *L’archeologia del sapere*; Foucault, M. (1976). *Histoire de la sexualité: Vol I—La Volonté de savoir*. Gallimard; tr. It. *Storia della sessualità, vol. I – La volontà di sapere*.

La produzione macchinica di documenti in combinato disposto con la produzione macchinica di realtà fa saltare l'idea che i fatti della scienza siano *fatti* – nel senso di *costruiti* o *instaurati* per usare il lessico latouriano – diversamente dai fatti sociali. E questo ha comportato quello che Paolucci ha definito un ritorno all'ideale medievale di verità, secondo il quale la verità risiede nelle persone, “a truth that is trust”. Accade infatti che da un lato la fiducia viene riposta nelle figure esperte che mediano il nostro accesso ai documenti che sono ampiamente inaccessibili; dall'altro accade che “when documents proliferate, a quantitative change becomes qualitative and truth as trust takes on the form of an anecdotal fact” (Ibidem). L'effetto quindi della produzione macchinica di documenti è un primato dell'esperienza sulla conoscenza, o, citando Ferraris ⁸³, “una privatizzazione dell'illusione di avere ragione”.

Proviamo a trarre due ordini di problemi da quanto detto sinora, nella convinzione che occorra riflettere ancora molto su questo.

- i) Partiamo dal primo, cioè da quella che Caroline Criado Perez nel suo libro “Invisibili. Come il nostro mondo ignora le donne in ogni campo. Dati alla mano”⁸⁴ chiama “la prevalenza del maschile-ove-non-altrimenti-indicato”, un problema che investe la storia dell'umanità sin dalla teoria dell'Uomo cacciatore, con un “enorme vuoto di dati” o “gender data gap”. Ora nell'economia del nostro discorso non c'è spazio per una trattazione specifica su questa “presenza-assenza” che semioticamente potremmo definire “la presentificazione di un'assenza” per gli effetti di senso che produce; tuttavia, ci sembra centrale nel valutare con più attenzione un *tentativo di salvataggio o di riscatto del sapere scientifico*, ancor di più considerando la mediazione macchinica nella produzione dei Big Data come *fatti di scienza*. Se la validità di una conoscenza non si basa sulla verifica di un'esperienza individuale costruita come fatto anedddotico, ma sull'universalizzazione di un modello teorico e sul principio di falsificabilità, allora l'intero libro di Perez è una messa in discussione per via di falsificazione del “canone” disciplinare di un certo numero di domini del sapere contemporaneo, dall'archeologia alla medicina. Il pregio di questa operazione, ci sembra, è che ciò non venga fatto - a beneficio di scienza - questionando “se il produttore di questo o quell'attrezzo a misura di maschio sia o non sia un sessista *in pectore*: le motivazioni personali sono irrilevanti, perché ciò che conta davvero è lo schema di fondo. [...] La mia tesi è che il vuoto dei dati di genere sia al tempo stesso causa

⁸³ Ferraris, M. (2021). *Documanità* [Documanity]. Laterza, p. 21

⁸⁴ Perez, C. C., (2020). *Invisibili: Come il nostro mondo ignora le donne in ogni campo. Dati alla mano*. Torino: Einaudi.

ed effetto di quella sorta di *non-pensiero* che concepisce l'umanità come quasi soltanto maschile". Poi continua riallacciando il filo del nostro ragionamento sui sistemi di AI: "Vi mostrerò che persino nel nostro mondo iperrazionale guidato da inflessibili supercomputer le donne sono ancora in buona parte il «secondo sesso» di cui parlava Simone de Beauvoir, un secondo sesso che rischia di essere trattato – quando va bene – come una sottocategoria dell'umano." (Ivi, p. 6). Nell'impalcatura concettuale di Perez i dati e le informazioni provengono da molte fonti, comprese le statistiche e le esperienze umane che sono sempre *situate*⁸⁵. Così Donna Haraway descriveva il concetto di sapere situato: "Volevamo andare oltre la rivelazione dei pregiudizi della scienza (era comunque troppo facile) e fare di più che separare i buoni scienziati dai cattivi pregiudizi e abusi. Sembrava promettente farlo attraverso strategie costruzioniste che non consentivano di ridurre la questione a polarizzazioni come pregiudizio contro oggettività, uso contro abuso, scienza contro pseudoscienza. Abbiamo smascherato le dottrine dell'oggettività perché minacciavano il nostro senso nascente di una soggettività e responsabilità storica collettiva, di nostri resoconti della verità 'radicati nel corpo', e siamo finiti con una scusa in più per non imparare la fisica del dopo-Newton, e con una ragione in più per abbandonare la vecchia pratica femminista fai-da-te di accomodarci da sole la macchina. Se sono solo testi, ridateli ai maschi." (Ivi, p. 107). Da una prospettiva semiotica, o almeno di una sua articolazione testualista, la lapidaria dichiarazione "Se sono solo testi, ridateli ai maschi" sembra condurre al riconoscimento di uno scacco, un'impossibilità di fondo a fare di questa semiotica una scienza posizionata, sul modello di altre scienze umane e sociali.⁸⁶

⁸⁵ Haraway, D. (1988). *Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective*. *Feminist Studies*, 14(3), 575–599; tr. It. *Saperi situati. la questione della scienza nel femminismo e il privilegio di una prospettiva parziale* in *Manifesto Cyborg. Donne, tecnologie e biopolitiche del corpo*. Feltrinelli (1995).

⁸⁶ Lorusso, M. M. (2025) nella sua tesi di dottorato "*Reading distress as oppression and resilience as resistance: a structural ecological analysis of trans and nonbinary people's experiences in Italy*" soppesa gli effetti della pratica del *posizionarsi* di chi fa ricerca soprattutto in relazione allo studio del vissuto esperienziale di individui appartenenti a "comunità storicamente marginalizzate". Citiamo: "This practice aims to challenge the illusion of neutrality traditionally maintained in dominant psychological research, which often reflects the perspectives of hegemonic societal groups. [...] Sharing one's positionality in research does not merely involve listing one's identity backgrounds without explaining how these identities influence the research process. Therefore, positionality should not be regarded merely as filler language, nor as a simple enumeration of identities without elucidating their integration into the research process, from design through recruitment, analysis, and writing, without engaged critical reflection on their embodied positionality" (Ivi, p. 30). Per dare corpo e sostanza a questa postura di ricerca, Lorusso convoca la *reflexive practice*: "Rather than asking, 'How am I influencing the data?', which frames embodiment as a bias to be minimized, researchers adopting a reflexive lens recognize their background and positioning not as threats to validity, but as integral elements of the research process" (Ivi, p. 31). Laddove però oggetto di ricerca

Riteniamo tuttavia che i testi, quando adeguatamente chiusi, possano essere più che sufficienti a dirci cosa ci sia al loro esterno, cosa si muova intorno ai loro confini o bordi.⁸⁷ D'altra parte, quello che Haraway qui denuncia è “una specie di terapia di elettroshock epistemologico”, nella misura in cui il bisogno/desiderio di sviluppare uno strumento di ricerca che consentisse di decostruire “le pretese di verità di una scienza ostile”, si è trasformato in una sottrazione dai tavoli da gioco dove si decidono “le verità pubbliche”. Non riteniamo che la semiotica sia una “scienza ostile” poiché tentativamente osserva, ricerca e questiona i rapporti tra le parti, senza pretesa di ricostruzione di un tutto. sosteniamo invece l'ipotesi che ribadiremo in conclusione, che La profondità della posizione di Haraway ci sembra ancora estremamente attuale, soprattutto nella messa a fuoco di una tensione epistemologica nella quale si tengono insieme una molteplicità di bisogni, ossia quel desiderio molteplice che anima il progetto di una “successor science” (Harding 1986)⁸⁸. Dice Haraway: “Così, penso che il problema mio e ‘nostro’ sia come ottenere simultaneamente una spiegazione di radicale contingenza storica per tutti i sistemi e soggetti di conoscenza, una pratica critica utile a riconoscere le ‘tecnologie semiotiche’ con le quali creiamo significati, e anche un impegno rigoroso, volto a ottenere resoconti fedeli di un mondo ‘reale’, un mondo che può essere parzialmente condiviso, e sia aperto a progetti mondiali di libertà circoscritta e adeguata abbondanza materiale, che portino un modico di sofferenza e un po' di felicità” (Haraway 1988, p. 109). Evidentemente non è l'utopia di un mondo perfetto o del migliore dei mondi possibili a muovere questa analisi. Siamo di fronte ad una disamina che, di fronte alla crisi del

non sono i vissuti soggettivi e le esperienze individuali, ma la vita dei segni e il loro modo di significare, come è possibile rendere conto dell'influenza che le identità esercitano sul fare ricerca senza travalicare soglie disciplinari? Lasciamo aperto l'interrogativo poiché meriterebbe un discorso sul metodo e sulla definizione del campo di ricerca, che per noi sono i luoghi dove riteniamo sia possibile declinare un discorso sulla riflessività (Cfr. Levit H. M. 2025; 2024). In questo senso sarebbe da ripartire dall'approccio material-discorsivo di Barad che abbiamo introdotto nelle pagine precedent.

⁸⁷ Riprendiamo da Dondero (2020): “Un enunciato si caratterizza per il fatto che è stato prodotto a partire da un atto che gli ha assicurato i limiti e le frontiere (la cornice, nel caso del dipinto; la copertina e la quarta di copertina nel caso del romanzo, i titoli di testa e i titoli di coda nel caso del film, ecc.)” (Ivi, p. 35). E sempre Dondero sulla cornice: “Nella pittura, è grazie alla chiusura assicurata dalla cornice che è possibile postulare che il dipinto funziona come un microsistema linguistico che permette di identificare i centri di attenzione, le relazioni tra il centro e la periferia, gli equilibri tra i differenti piani di profondità, l'economia delle forze di direzione della luce, ecc...” (Ivi, p. 69 nota 48).

⁸⁸ Harding, S. (1986). *The science question in feminism*. Cornell University Press. Nel capitolo *From feminist empiricism to feminist standpoint*, Harding passa in rassegna cinque ragioni che giustificano un'epistemologia femminista e posizionata come ancorata “in a theory of gendered activity and social experience” (Ivi, p. 141).

moderno e della scienza che l'ha prodotto, ne deriva tutte le conseguenze in termini di "irreducible difference and radical multiplicity of local knowledges".

Il desiderio molteplice è pericoloso, apre al paradosso e necessariamente contraddittorio; essere femminista e fare scienza, per Haraway, non significa andare alla ricerca di una "dottrina dell'oggettività che promette trascendenza, né di un racconto che perde il filo delle sue mediazioni proprio nel punto in cui si può essere riconosciuti responsabili di qualcosa, né di illimitato potere di strumentalizzazione. Non vogliamo che sia una teoria di poteri innocenti a rappresentare il mondo, dove corpi e linguaggi cadono nel gaudio della simbiosi organica, né vogliamo teorizzare il mondo, e ancor meno agire al suo interno in termini di Sistemi globali. Eppure, abbiamo bisogno di una rete di rapporti che copra il mondo, e includa l'abilità di tradurre parzialmente conoscenze tra comunità molto differenti e differenziate in termini di potere. Abbiamo bisogno del potere delle moderne teorie critiche su come significati e corpi vengono costruiti, non a scopo di negare significati e corpi, ma per costruire significati e corpi che abbiano un futuro" (Ivi, p. 109).

- ii) A questo livello si collega il secondo ordine di problemi, che porta a domandarci se la fame di verità od onestà algoritmica, su cui si basa la *truthful AI*, non sia in qualche modo un tentativo di invertire la tendenza verso una ricentralizzazione dei rapporti di sapere e potere cristallizzati in una supposta oggettività resa possibile da un modello di intelligenza universale che crede solo a se stessa. Questo "se stessa" sta per un archivio organizzato, programmato e programmabile da esseri umani che, pur supponendo l'assoluta buona fede, non vedono o non possono vedere la mancanza; la stessa impossibilità di cui parlava Harding.

Per provare a comprendere meglio qual è la natura dell'output prodotto dalla macchina, dobbiamo fare un passo indietro o meglio all'interno dei processi e delle strutture alla base della generazione degli enunciati macchinici.

2.0 Per una definizione dei Large Language Models

Articolare storicamente il passato non significa
conoscerlo «come propriamente è stato».
Significa impadronirsi di un ricordo come
esso balena nell'istante di un pericolo

(Benjamin, W. *Tesi di filosofia della storia*).

Non è che il passato getti la sua luce sul presente
o il presente la sua luce sul passato,
ma immagine è ciò in cui quel che è stato
si unisce fulmineamente con l'ora in una costellazione.
In altre parole: immagine è la dialettica nell'immobilità.
Poiché, mentre la relazione del presente con il passato
è puramente temporale, continua, la relazione
tra ciò che è stato e l'ora è dialettica:
non è un decorso ma un'immagine discontinua, a salti.
Solo le immagini dialettiche sono autentiche immagini
(cioè non arcaiche); e il luogo, in cui le si incontra,
è il linguaggio.

(Benjamin, W. *Appunti e materiali*).

Proveremo quindi a definire ordinatamente il funzionamento di un Large Language Model; per fare questo ricorreremo a una letteratura in parte specialistica, proveniente da settori scientifici come la computer science e la matematica, in parte divulgativa per consentire di mantenere “il filo del discorso” che spesso nei paper e nella manualistica tecnica si perde. Il nostro obiettivo non è quello di sostituirci a studiosi e studiosi di settore, ma costruire ponti tra ambiti disciplinari per favorire uno scambio di concetti e strumenti, altrimenti inconciliabili. Ricordiamo in proposito che nel 1959 C. P. Snow⁸⁹ vedeva lo iato tra *scientists* e *literary intellectuals* (assimilati alla cultura tradizionale) spaccare la vita intellettuale della società occidentale post Seconda Guerra Mondiale in *due culture*. Quale fu l'effetto di questa scissione? Così lo descrive Snow:

“This polarisation is sheer loss to us all. To us as people, and to our society. It is at the same time practical and intellectual and creative loss [...] The degree of incomprehension on both sides is the kind of joke which has gone sour.” (Ivi, p. 12).

E più avanti, dopo aver analizzato le mancanze reciproche e definito i rispettivi valori epistemici continua “The clashing point of two subjects, two disciplines, two cultures – of two galaxies, so far

⁸⁹ Snow, C. P. (1959), The two cultures and the scientific revolution, in *The Rede Lecture 1961*, Cambridge University Press.

as that goes – ought to produce creative chances. In the history of mental activity that has been where some of the breakthroughs came. The chances are there now. But they are there, as it were, in a vacuum, because those in the two cultures can't talk each other" (ivi, p. 17).

Possiamo dire che sia così ancora oggi? Nel seguito analizzeremo i discorsi di alcuni domini disciplinari direttamente coinvolti nel costruire narrazioni intorno all'Intelligenza Artificiale, se non direttamente coinvolti nello sviluppo di sistemi di apprendimento automatico nelle sue declinazioni più recenti. Anticipiamo che i bacini di produzione discorsiva ci sembrano risentire fortemente di posture epistemologiche che segnalano differenze precipue tra campi disciplinari. Si levano scudi in difesa e in attacco su cosa sia o non sia una macchina intelligente, quale stadio di sviluppo tecnologico è stato raggiunto e quali conseguenze ed effetti tutto ciò comporterà per l'umanità come specie. In parte il vuoto di cui parlava Snow appare saturo di prese di parola, costruendo un ventaglio di interpretazioni dal profilo mitologico, dove cioè coesiste tutto e il contrario di tutto.

Liu et al. (2025)⁹⁰ definiscono la modellazione del linguaggio (Language Modelling – LM) come un approccio fondamentale per ottenere un'intelligenza cognitiva nel campo del Natural Language Processing; per intelligenza cognitiva intendono la capacità di comprendere, generare e manipolare il linguaggio umano per il funzionamento di sistemi tecnici quali traduzione automatica, chatbots, sentiment analysis e riassunti. Grazie all'evoluzione del deep learning⁹¹ i primi modelli di linguaggio su base statistica (Statistical Language Model – SLM) si sono trasformati nei modelli basati su reti neurali (Neural Language Models – NLM). A fondamento di questa evoluzione vi è il ricorso a tecniche come il *word embedding* che consente di rappresentare la parola in uno spazio n-

⁹⁰ Liu, Y., He, H., Han, T., Zhang, X., Liu, M., Tian, J., Zhang, Y., Wang, J., Gao, X., Zhong, T., Pan, Y., Xu, S., Wu, Z., Liu, Z., Zhang, X., Zhang, S., Hu, X., Zhang, T., Qiang, N., ... Ge, B. (2025). Understanding LLMs: A comprehensive overview from training to inference. *Neurocomputing*, 620, 129190. <https://doi.org/10.1016/j.neucom.2024.129190>

⁹¹ A questo livello del discorso basti sapere che la storia del deep learning non è così recente come si può supporre dall'esplosione del termine associata alla diffusione dei Large Language Model e in particolare dalla messa in circolazione per il grande pubblico del bot ChatGPT di OpenAI a partire da novembre 2022. Goodfellow et al. (2017) fanno risalire la prima ondata di sviluppo del deep learning, sotto altre vesti, alla teoria cibernetica degli anni '40-'60 del Novecento. A questa prima ondata ne sono seguite due successive, una conosciuta come connessionismo tra gli anni '80 e '90, l'altra effettivamente denominata *deep learning* che ha preso avvio nel 2006. Quando Ian Goodfellow, Yoshua Bengio e Aaron Courville pubblicarono uno dei manuali più noti sull'insieme di tecniche di machine learning conosciute come deep learning, ancora non era esploso il fenomeno ChatGPT. Era infatti il 2017. Oggi ci si potrebbe domandare rispetto alla periodizzazione da loro proposta, se l'avvento dei Large Language Model rappresenti la lunga coda della terza onda dello sviluppo del deep learning oppure se abbia più senso parlare di un cambio di paradigma all'interno del campo largo degli studi sull'Intelligenza Artificiale. Riteniamo in ogni caso non sia questa la sede in cui affrontare una tale questione.

dimensionale attraverso un vettore distribuito. Nell'apprendimento non supervisionato all'interno del Natural Language Processing, l'applicazione del *word embedding* ha la funzione di rappresentare testi non etichettati. Prevede l'utilizzo di vari metodi – Variational AutoEncoders (VAEs) e Generative Adversarial Networks (GANs) – che creano spazi latenti per la modellazione e la sintesi dei dati. Come osservano Liu et al. (2019)⁹², la rilevanza degli spazi latenti ha natura duplice: da un lato consentono di “vedere” relazioni tra i dati di cui non si ha consapevolezza, dall'altro possono servire da spazi di funzione per applicazioni di machine learning supervisionato a valle. Questi stessi principi di funzionamento sono applicabili nello specifico all'embedding layer di un Transformer, dove la rappresentazione latente che si costruisce in fase di addestramento sembra inerire meno una conoscenza linguistica generica e più le caratteristiche geometriche dello spazio latente prodotto dal Transformer (Marbut et al. 2024).⁹³ In particolare, descrivono gli spazi latenti come altamente anisotropici, con i dati concentrati in un “cono” ristretto e non distribuiti in tutto lo spazio e le dimensioni a disposizione. Tuttavia, l'incremento dell'isotropia generale non è vantaggioso nei modelli con spazio latente contestuale. In questi casi sembra funzionare meglio aumentare l'isotropia nei clusters locali per il miglioramento della performance del modello.⁹⁴ Altro metodo studiato per

⁹² Liu, Y., Jun, E., Li, Q., & Heer, J. (2019). Latent Space Cartography: Visual Analysis of Vector Space Embeddings. *Computer Graphics Forum*, 38(3), 67–78. <https://doi.org/10.1111/cgf.13672>

⁹³ Marbut, A.C., Chandler, J.W., & Wheeler, T.J. (2024). *Exploring the Impact of a Transformer's Latent Space Geometry on Downstream Task Performance*. ArXiv, abs/2406.12159. “We propose that at least some of the benchmarking performance seen with this pre-training/finetuning protocol can be attributed to geometric characteristics of the latent space that result from the transformer architecture, divorced from any generic linguistic knowledge. Many common downstream tasks for language models, such as classification or paraphrase, use measures of similarity or separability in their loss functions. These measures rely on the way that latent representations are organized in this space, so it stands to reason that there could be a measurable characteristic of this organization that would be advantageous for learning these fine-tuning tasks. If the latent space is already organized with data placed in various high dimensional nooks and crannies, it may be easier for a powerful transformer model to learn how to access the data appropriately during fine-tuning, even if the organization scheme is non-linguistic” (Ivi, p. 1). Il nucleo della tesi sostenuta è che a fronte dell'efficacia dei modelli di linguaggio basati sul transformer e sul paradigma pre-training/fine-tuning, la ragione dell'impressionante capacità di questi modelli nei test di benchmarking non risieda in un pre-addestramento su task linguistici e dunque su una conoscenza linguistica generalizzata. Sono le caratteristiche geometriche in sé dello spazio latente a rendere questi modelli performanti nei test di benchmarking.

⁹⁴ Cfr. il paragrafo successivo 3.4 per le tecniche di training basate su metaeuristiche nell'esplorazione dei *search spaces*; anche in questo studio emerge come il comportamento del vivente possa fungere da esempio nello studio della distribuzione spaziale dei dati nello spazio latente. Viene infatti utilizzata la metafora ecologica della “mean crowding” di Lloyd, cioè la densità vista dagli occhi di un singolo individuo, il cui posto nello spazio latente verrebbe occupato da centroidi. Si immagina di avere un insieme di dati proiettato in uno “spazio latente” e di averlo quantizzato in celle/cluster attorno a dei centroidi (ogni punto viene assegnato al centroide più vicino). Se i dati sono distribuiti in modo davvero uniforme, ogni cella dovrebbe contenere più o meno lo stesso numero di punti. Se invece alcune celle sono molto affollate e altre quasi vuote, la distribuzione è irregolare o

migliorare la performance nel benchmarking è la quantizzazione, ossia un metodo di compressione nel quale dati ad alta-dimensionalità sono clusterizzati in sottospazi che possono essere ricombinati per recuperare la dimensionalità originaria con una perdita minima. Con questo metodo lo spazio latente diviene più facilmente separabile e dunque favorire una preparazione dei dati migliore per la parte di fine-tuning sui task a valle.

La difficoltà di studio degli spazi di latenza, soprattutto in relazione alla distribuzione della rappresentazione delle parole nella lingua naturale nei modelli di linguaggio contestuali, nasce dal fatto che la forma della rappresentazione dei token nello spazio latente è fortemente influenzata dal contesto in cui i token vengono “visti”; dunque, ciascun token, può avere un numero infinito di rappresentazioni. In 4.0 vedremo come questa è l’indeterminazione determinabile che caratterizza l’enciclopedia di Eco nella sua globalità e vedremo anche l’analogia geometrica che Eco costruisce a riguardo.⁹⁵

Un’ulteriore evoluzione in termini di framework di sviluppo è rappresentata dai Pre-trained Language Model (PLM), di cui la famiglia di modelli GPT fa parte (sappiamo infatti che l’acronimo GPT sta per Generative Pre-Trained Model). Prima dei modelli generativi basati sull’architettura Transformer, di cui parleremo tra un momento, sono stati fatti altri esperimenti tra i cosiddetti modelli pre-trained. Gli autori e le autrici citano come esempi, implementato nel 2018 (Peters et al. 2018)⁹⁶,

“a chiazze” (patchy). Da qui Marbur et al. costruiscono un indice di misurazione della densità dei dati nello spazio latente.

⁹⁵ Per una trattazione semiotica del concetto di spazio latente si vedano Cristina Voto (2023), *Verso una semiotica della spazialità latente*, in Leone, M. (2023b), *Il volto latente*, FACETS Digital Press, Open Access, dove Voto propende per una lettura simulativa, per quanto accurata, del rapporto tra enunciazione spaziale e generatività dello spazio di latenza. Dice: “Allenare un modello computazionale a imparare le somiglianze significa, infatti, insegnargli a classificare, quantitativamente i dati imparando le caratteristiche. Ma questo apprendimento richiede un passaggio fondamentale, una semplificazione degli stessi dati affinché sia possibile trovare degli schemi, dei pattern. Man mano che la rete neurale apprende impara le caratteristiche dei dati e mappa, simulando, gli schemi appresi attraverso il riconoscimento di caratteristiche similari. Ma di che tipo di similitudine stiamo parlando? Una similitudine che potremmo pensare in termini topologici e quantitativi piuttosto che iconici e qualitativi, per questo credo che piuttosto che un’enunciazione spaziale ci troviamo di fronte a una sua simulazione” (Ivi, pp. 226-227). Ancora, si veda la prima sezione del numero di *Signata*, curata D’Armenio, E., Rosso, A., & Voto, C. (2025), *Sémiotique de l’espace à l’ère de la logique computationnelle et des pratiques numériques*. *Signata*. *Annales des sémiotiques / Annals of Semiotics*, 16, Articolo 16. <https://doi.org/10.4000/14rhb>, dove Bruno Bachimont (2025) e Alban Leveau-Vallier (2025) elaborano rispettivamente secondo una prospettiva struttural-generativa e una logica associativa, la produzione di senso e l’organizzazione del sapere all’interno dello spazio latente dei modelli generativi.

⁹⁶ Peters M. E., et al. (2018), “Deep contextualized word representations,” in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 1 (Long Papers), Jun., pp. 2227–2237

che introduceva ELMo (*Embeddings from Language Models*) una “deep contextualized words representation” capace di modellizzare:

- i) feature complesse delle parole utilizzate (sintassi e semantica);
- ii) come questi usi varino tra i diversi contesti linguistici (polisemia).

Il tipo di embedding utilizzato in questo sistema derivava tuttavia da un’architettura Long-Short-Term-Memory (LSTM) bidirezionale e non dal più recente Transformer, che sfrutta invece i meccanismi di auto-attenzione in parallelo.

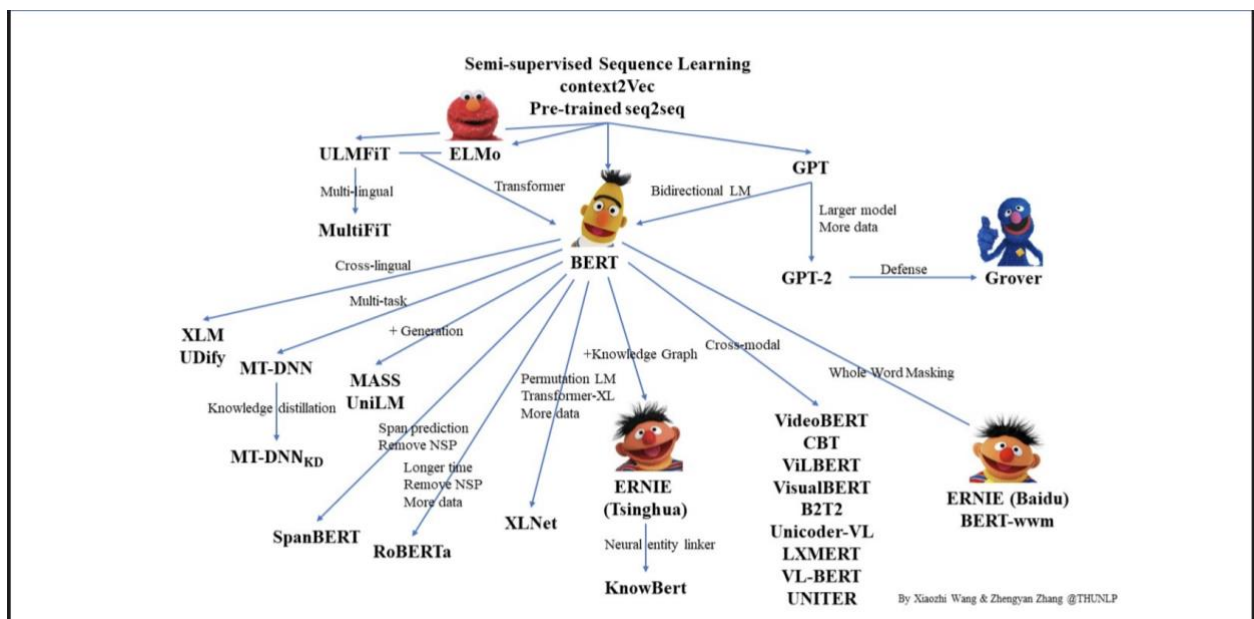


Figura 8: Sintesi visiva dell'evoluzione dei modelli Transformer⁹⁷

La caratteristica specifica di questa architettura è di essersi prestata al cosiddetto scaling-up dei modelli, ossia la crescita di scala del modello di linguaggio, per l'appunto large, dovuta sia a un incremento dell'ordine dei terabyte nei dataset di addestramento che alla dimensione misurata sul numero di parametri. I modelli attuali arrivano a sfiorare le migliaia di miliardi, a riconferma di quella che è stata definita “the scaling hypothesis”⁹⁸.

⁹⁷ Chersoni, EM (2021), *Distributional Semantics and Event Knowledge*: https://www.uni-hildesheim.de/media/fb3/informationwissenschaft/HIER/Chersoni_GEK_Presentation.pdf [ultimo accesso: 2 ottobre 2025]

⁹⁸ Per una ricostruzione di questa evoluzione di scala dei modelli di linguaggio si veda: Cristiani, N. (2025), *Sovrumano. Oltre i limiti della nostra intelligenza*, Il Mulino. Qui possiamo leggere: “I sistemi di IA stanno rapidamente cambiando ordine di grandezza, e questo cambiamento di scala è la chiave del loro successo. Ma gli strumenti culturali e tecnici per comprenderli stanno facendo fatica a seguirli, perché sono stati sviluppati a livelli diversi. La rete neurale LeNet negli anni Novanta aveva circa 60 mila parametri; AlexNet nel 2012 aveva circa 60 milioni di parametri; GPT 2 aveva circa 1

Occorre però sottolineare come non vi sia consenso unanime all'interno della comunità di chi fa ricerca e implementa tecniche di IA relativamente all'ipotesi dello scaling-up dei modelli. Difatti vi è una componente consistente di esponenti della cosiddetta IA simbolica che rimarca l'inadeguatezza della sola crescita dimensionale dei modelli sia dalla prospettiva dell'emergenza di capacità non specificatamente previste o programmate, sia – ma anche di conseguenza – in termini di *explainability* degli stessi modelli e dei loro output: ci riferiamo a titolo di esempio ai lavori di Gary Marcus e di John Sowa presentati all'Ontology Summit 2024, rispettivamente intitolati “*No AGI (and no Trustworthy AI) without Neurosymbolic AI*” e “*Without Ontology, LLMs are clueless*”.⁹⁹

Così Cristianini introduce l'ingresso in scena del Transformer: “La notizia in quei giorni era AlphaZero, il nuovo algoritmo di DeepMind che aveva imparato interamente da solo a giocare a Go e a scacchi meglio di ogni altro algoritmo, allenandosi contro sé stesso per 40 giorni, usando centinaia di processori [...] AlphaZero catturò così tanto la nostra attenzione in quei giorni che non prestammo attenzione a un articolo pubblicato in quella stessa conferenza da alcuni ricercatori di Google che descriveva un nuovo metodo per risolvere compiti di traduzione molto rapidamente”. (Cristianini 2024, p. 25).

La conferenza a cui Cristianini fa riferimento è la *Neural Information Processing Systems* (NeurIPS), che ricorre annualmente dal 1987 e che proprio nel 2017, alla sua trentunesima edizione, vide la presentazione alla comunità scientifica del Transformer. Nelle parole di Cristianini, un transformer “ricombinando in modo diverso una serie di meccanismi già noti, consentiva di esaminare parti diverse della frase indipendentemente le une dalle altre, e quindi simultaneamente. Uno di questi meccanismi, che i tecnici chiamano «attenzione», consentiva all'algoritmo di scoprire quali altre parole devono essere considerate quando si traduce una parola data, ovvero da quali parole dipende la sua interpretazione. Non solo l'algoritmo poteva scoprire le «dipendenze», ma poteva imparare a farlo automaticamente, con l'esperienza, e lo faceva molto rapidamente” (Ivi, pag. 26).

miliardo di parametri; GPT 4 circa 1.000 miliardi. Queste dimensioni non solo richiedono infrastrutture e competenze diverse, ma danno vita a fenomeni diversi.”

⁹⁹ Per visualizzare slide e video registrazione degli interventi visitare accedere al link: <https://ontologforum.com/index.php/OntologySummit2024> [ultimo accesso: 29 settembre 2025].

Sotto il profilo tecnico, il team di Google in un lavoro congiunto i cui meriti sono dettagliatamente attribuiti e distribuiti nel paper “*Attention is all you need*”¹⁰⁰, spiega che questa nuova architettura di rete sostituisce ai *recurrent* e *convolutional layer* normalmente utilizzati nei modelli encoder-decoder, il primo modello di transduzione sequenziale interamente basato su un meccanismo *multi-headed self-attention*.

Quando si parla di architettura, si intende la struttura generale della rete (Goodfellow et al. 2017, p. 197)¹⁰¹, ossia quante unità o neuroni artificiali dovrebbe contenere e come queste unità sono collegate tra loro. La maggior parte delle reti sono organizzate in gruppi di unità definiti layer, a loro volta costituenti delle strutture a catena in cui ciascun layer è funzione del layer che lo precede. Ciò che occorre definire, quindi scegliere, è la profondità del network (numero di strati/layer) e la larghezza di ciascun layer (numero di unità/neuroni artificiali). Sembra non esserci una regola che definisca la dimensione, ma l’architettura ideale è il risultato di una sperimentazione guidata dal monitoraggio dell’errore dell’insieme di convalida.

Altra caratteristica che rende particolarmente efficaci questi modelli a confrontarsi con compiti complessi implicanti il processamento del linguaggio naturale è che si prestano al cosiddetto *parallel computing*, ossia a svolgere in parallelo numerosi processi computazionali consentendo quindi interazioni indipendenti tra i token in input, successivamente mediate tramite prodotto scalare interno tra gli embedding di query e di key-value contenuti in ciascuna unità/testa del meccanismo di auto-attenzione (Sandford et al. 2024)¹⁰².

Goodfellow et al. (2017) spiegano che le reti neurali profonde vengono solitamente definite come *Feed-Forward Neural Network* (FFN) oppure *MultiLayer Perceptrons* (MLP). Il percettrone è la rete neurale più semplice che non contiene layer nascosti (hidden). Una Feedforward Neural Network ha come obiettivo di funzionamento quello di approssimare una funzione f ottimale (in gergo tecnico presenta questa notazione f^*). Riportano come esempio una classificazione tipo 0-1, dove la funzione $y = f^*(x)$ mappa un input x a una categoria y ; la rete neurale definisce una mappatura $y = f(x; \theta)$, dove θ sono i parametri che la rete apprende o, più precisamente, i valori assunti dai parametri e che meglio approssimano la funzione.

¹⁰⁰ Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (No. arXiv:1706.03762). arXiv. <http://arxiv.org/abs/1706.03762>

¹⁰¹ Goodfellow, I., Bengio, Y., & Courville, A. (2017). *Deep Learning*. MIT Press.

¹⁰² Sanford, C., Hsu, D., & Telgarsky, M. (2024). *Transformers, parallel computation, and logarithmic depth* (No. arXiv:2402.09268). arXiv. <https://doi.org/10.48550/arXiv.2402.09268>

Si chiamano reti *feedforward* – feedforward è raramente tradotto in italiano – poiché i flussi informativi scorrono in avanti, spesso senza retroazione (in caso di feedback si parla di recurrent neural network); nel dettaglio l'informazione attraversa la funzione che viene elaborata a partire da x per ottenere l'output y , passando per le computazioni intermedie finalizzate a definire la funzione f .

Tali reti sono organizzate in network, cioè compongono, secondo un certo ordine che è associato a un grafo aciclico orientato/diretto, funzioni differenti connesse in una catena¹⁰³. La lunghezza complessiva della struttura a catena indica la profondità del modello.

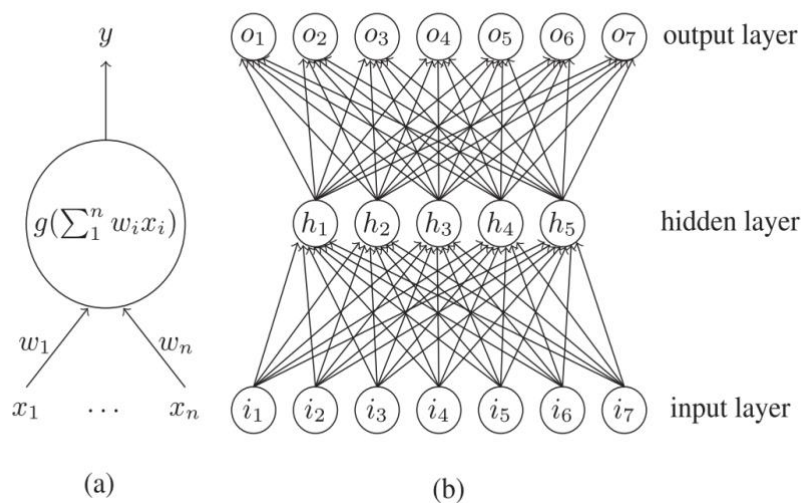


Figura 9: a) Neurone artificiale; b) Feed-forward neural network con un hidden layer¹⁰⁴

L'architettura Transformer è costituita di due moduli principali, un Encoder e un Decoder, all'interno dei quali è presente il meccanismo di attenzione (Self-Attention e Multihead Attention)¹⁰⁵. Partiamo dal meccanismo di attenzione per poi tornare ai primi due.

¹⁰³ L'esempio riportato in Goodfellow et al 2017 è: ci sono 3 funzioni $f(1)$, $f(2)$ e $f(3)$, connesse in una catena a formare $f(x) = f(3)(f(2)(f(1)(x)))$. In questo esempio, $f(1)$ rappresenta il primo layer della rete, $f(2)$ il secondo layer ed $f(3)$ il terzo.

¹⁰⁴ In Lenci A, Sahlgren M. Neural Network Models. In: *Distributional Semantics. Studies in Natural Language Processing*. Cambridge University Press; 2023:144.

¹⁰⁵ Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (No. arXiv:1706.03762). arXiv. <http://arxiv.org/abs/1706.03762>

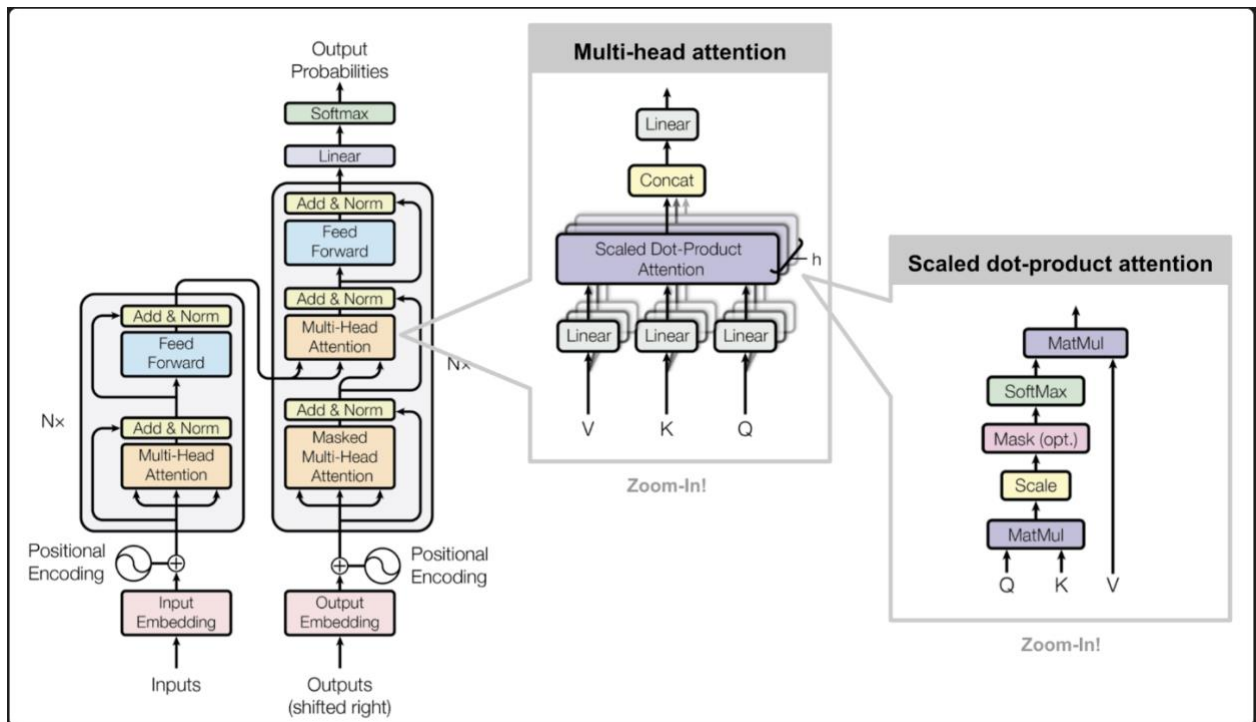


Figura 10: L'architettura modello completo del Transformer.¹⁰⁶

L'attenzione ha come obiettivo primario quello di selezionare una piccola quantità di informazioni importanti all'interno un grande quantitativo di dati, focalizzandosi su queste porzioni più rilevanti rispetto al resto. Una variante del meccanismo di attenzione è la self-attention che “reduces reliance on external information and excels at capturing internal correlations within data or features” (Liu et al. 2025, p.2). Applicare questa funzione a un testo significa calcolare – attraverso una somma ponderata dei valori di ciascuna parola di cui si compone una frase e dove il peso è determinato dalla rilevanza che ciascuna parola ha rispetto alla parola target - la mutua influenza delle parole per ovviare al problema delle dipendenze a lungo raggio: il modello pesa l'importanza di differenti parole in una frase nel momento in cui deve predire la parola successiva. Il meccanismo di self-attention si articola in tre passaggi, ossia l'individuazione dei vettori: query, key e value. Citiamo:

“The query vector represents the word being attended to, while the key vectors represent all words in the sentence. The value vectors store information associated with each word. The attention weights are computed by taking the dot product between the query and the key vectors, followed by a SoftMax operation to obtain a distribution over the words” (Ibidem).

¹⁰⁶ Image source: Fig 1 & 2 in Vaswani, et al., 2017, elaborata da Weng, L. *Attention? Attention!* <https://lilianweng.github.io/posts/2018-06-24-attention/>. Rispetto alla figura 9b, ciascun layer della rete neurale è costituito dai blocchi base MHSA e FFN che definiscono dei sottolivelli.

La mappatura completa dell'attenzione viene rappresentata con la seguente formula:

$$A(Q, K, V) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V.$$

dove, il prodotto scalare tra i vettori di query e i vettori key, nella forma di una matrice trasposta, viene mitigato dal fattore scalare $\frac{1}{\sqrt{d_k}}$ evitando un incremento di dimensionalità e mantenendo stabile la magnitudine del prodotto. Nelle Figure 11-13, riprese da Vig (2019)¹⁰⁷ sono mostrate tre diverse tipologie di visualizzazione del meccanismo di attenzione in un modello Transformer: la prima rimanda a come il meccanismo agisce a livello di singola testa; la seconda offre una visualizzazione complessiva del meccanismo di attenzione a livello del modello; infine, la terza illustra l'attenzione a livello di singolo neurone. L'articolo compara l'utilità dello strumento di visualizzazione presentato sia rispetto al modello BERT che rispetto al modello GPT-2; sottolineiamo questo aspetto a partire dalla differenza di integrazione dell'architettura transformer tra i due modelli. BERT utilizza solo il blocco Encoder, viceversa GPT-2 solo il blocco Decoder.

Le immagini inserite in questo elaborato fanno riferimento solo al modello BERT. Vale la pena ricordare che le immagini hanno a tutti gli effetti funzione di *scaffolding cognitivo* nell'illustrare come agisce il meccanismo di attenzione e tuttavia, come nel disclaimer “L'immagine sulla confezione ha il solo scopo di presentare il prodotto”, non corrisponde a *ciò che è* il meccanismo di attenzione. Questo per due ragioni principali: una di carattere operativo, il modello non ha accesso direttamente alle parole, che prima di essere trasformate in vettori o matrici di embedding, sono sottoposte a tokenizzazione. L'altra è di carattere generale e in parte l'abbiamo già esplicitata parlando di “taglio agenziale” come postulato material-semiotico.

Per chiarire ulteriormente il punto citiamo Leone (2023b): “Sarebbe interessante chiedere agli specialisti di semiotica dell'intelligenza artificiale quanti di loro abbiano visto una rete neurale. Molti sarebbero probabilmente in grado di descriverla e persino di spiegarne il funzionamento. Ma quanti l'avranno vista davvero? Se si cercano immagini di reti neurali su Google o su qualsiasi altro motore di ricerca, o anche se si cercano immagini di componenti specifici di reti neurali, il risultato è deludente. Ad esempio, se si tenta di visualizzare un percettone (o neurone McCulloch-Pitts) [...] si

¹⁰⁷ Vig, J. (2019). A Multiscale Visualization of Attention in the Transformer Model. In M. R. Costa-jussà & E. Alfonseca (A c. Di), *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 37–42). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-3007>

ottengono formule matematiche, diagrammi e, nel migliore dei casi, immagini storiche [...]. Sebbene la miniaturizzazione sia una delle caratteristiche principali dello sviluppo tecnologico digitale, nel caso dell'intelligenza artificiale tale rimpicciolimento implica non solo un restringimento, bensì anche la vera e propria invisibilità. In larga misura, l'intelligenza artificiale è il prodotto di macchine invisibili, che vengono visualizzate attraverso metafore visive – spesso impregnate di pregiudizi ideologici – o formule o diagrammi logico-matematici, oppure immagini d'epoca” (Leone 2023b, p.13). Ruolo di una semiotica dell'IA sarà proprio quello di ricostruire i livelli di mediazione tra visibile e invisibile. Dice ancora Leone: “La riflessione sull'invisibile, o meglio sulla dialettica tra visibilità e invisibilità, è fondamentale per la semiotica. Quando parliamo di segni, stiamo parlando di relazioni tra percettibilità e impercettibilità, tra presenza e assenza, tra il manifesto e il latente. Nel modello di segno di Peirce, essenzialmente neokantiano, l'oggetto è il latente intrinseco che viene manifestato solo parzialmente e in modo obliquo da un interpretante, il quale coglie un aspetto di questa latenza e la mette in relazione con un *representamen* che è invece presente ai sensi e quindi alla mente” (Ivi, p. 14).

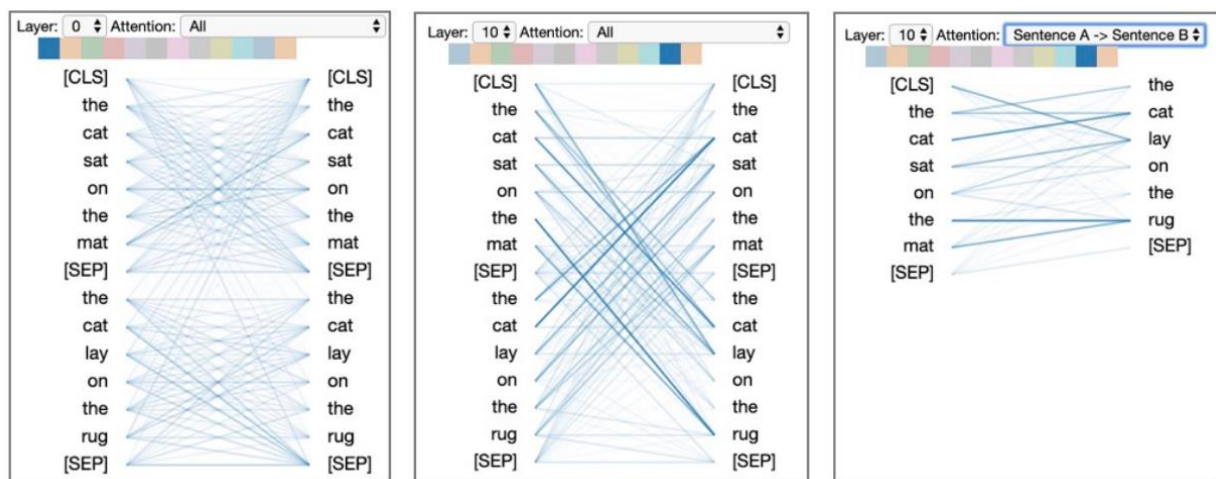


Figura 11: Attention-head view for BERT.¹⁰⁸

¹⁰⁸ In Vig 2019, p. 38: “for inputs the cat sat on the mat (Sentence A) and the cat lay on the rug (Sentence B). The left and center figures represent different layers / attention heads. The right figure depicts the same layer/head as the center figure, but with Sentence A → Sentence B filter selected”.

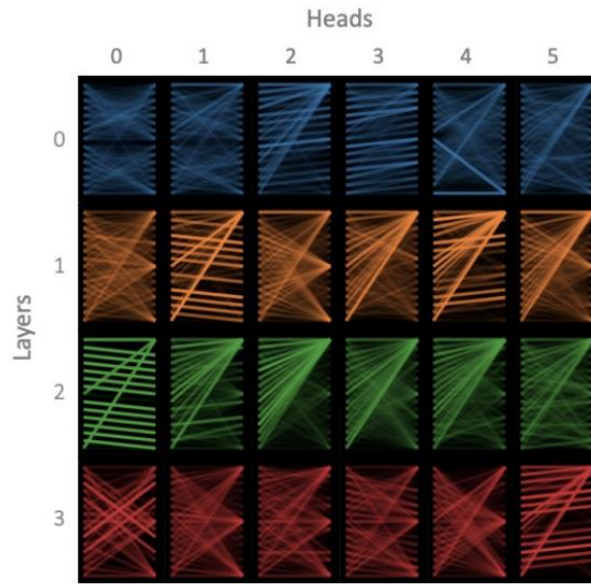


Figura 12: Model view of BERT, for same inputs as in Figure 11.¹⁰⁹

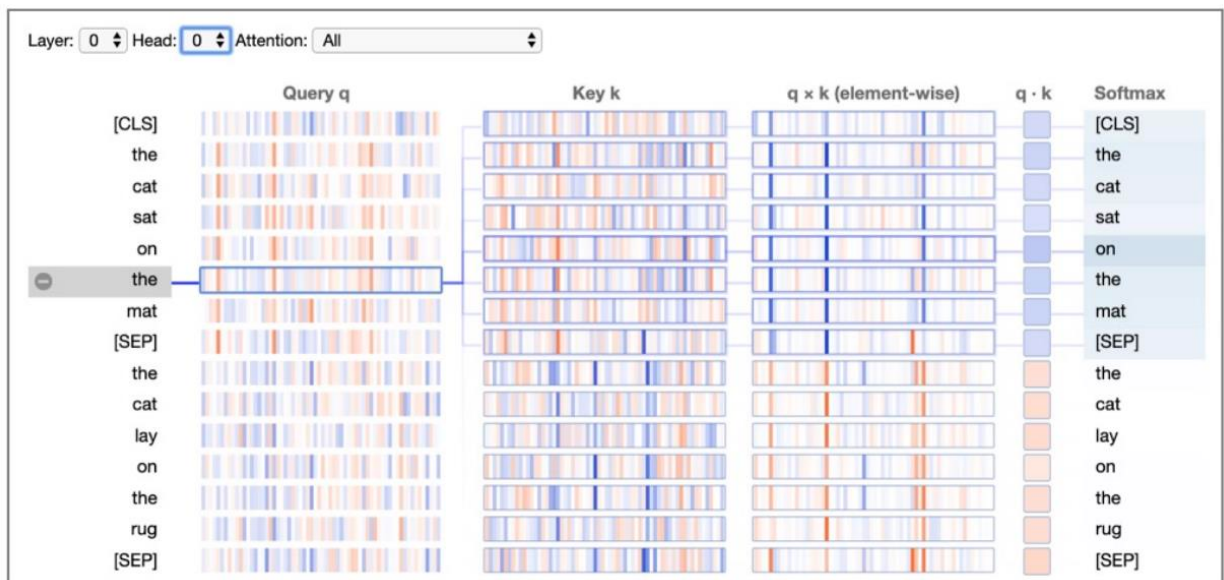


Figura 13: Neuron view of BERT for layer 0, head 0 (same one depicted in Figure 11, left).¹¹⁰

Torniamo all'invisibile quindi. Per quanto concerne il blocco Encoder, esso è composto di molteplici layer, ciascuno dei quali include un MultiHead Attention Mechanism (MHSA) e una Feed Forward Neural network (FFN). La sua funzione è di estrarre gradualmente le caratteristiche della sequenza di dati fornita in input; nello specifico, le dipendenze tra le diverse posizione nella sequenza di input.

¹⁰⁹ Excludes layers 4–11 and heads 6–11 (Vig 2019, p. 40).

¹¹⁰ Positive and negative values are colored blue and orange, respectively, with color saturation based on magnitude of the value. As with the attention-head view, connecting lines are weighted based on attention between the words (Vig 2019, p. 41).

La MHSA si concentra sulle singole posizioni, la FFN estrae ulteriori features dall'output del meccanismo di attenzione.

Il modulo Decoder presenta in via di principio le stesse caratteristiche del modulo Encoder, dunque un MHSA e un FFN, a cui si aggiunge un ulteriore livello di attenzione che ha lo scopo di nascondere (to mask) e limitare al modello l'accesso informazioni future man mano che predice il token successivo, mantenendo in questo modo un principio di causalità (matematicamente si tratta di probabilità condizionata) e prevenendo la perdita di dati.

Connesso alla funzione dell'Encoder, troviamo quello che viene definito *positional encoding* (Figura 10). Nel Transformer l'attenzione auto-guidata rende il modello potente ma, per costruzione, indifferente all'ordine della sequenza di input: senza ricorrenza né convoluzioni, la rete è *permutation-invariant* e quindi incapace, da sola, di distinguere “il gatto morde il cane” da “il cane morde il gatto”. Il *positional encoding* inietta nell'input in ingresso tracce di sequenzialità che consentono all'attenzione di pesare non solo che *cosa* sono i token, ma anche *dove* si trovano e *come* si relazionano lungo la frase. Le informazioni di posizione vengono rappresentate anch'esse con vettori di embedding posizionali che si sommano agli embedding delle parole (oppure influenzando direttamente i punteggi di attenzione). Così il modello distingue lo stesso token in posizioni diverse. Esistono diversi metodi di encoding posizionale, ma ciò che per noi rileva è che esso opera come un lettore di sequenze che permette di disambiguare somiglianze.

Definita per sommi capi l'architettura del Transformer, possiamo domandarci cosa accada a questa infrastruttura durante il training. Definiamo innanzitutto il training come il momento in cui un algoritmo di *machine learning* impara ad apprendere dai dati che sono forniti in fase di addestramento.

Proponiamo di seguito due definizioni, entrambe formulate all'interno di manuali di settore, con una differenza di soli 5 anni tra una pubblicazione e l'altra:

1. La prima è in Goodfellow et al. (2017, 99) a sua volta ripresa da Mitchell (1997): “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.”
2. La seconda è contenuta in Russel, S., Norvig P. (2022, 670): “An agent is learning if it improves its performance after making observations about the world... When the agent is a computer, we call it machine learning: a computer observes some data, builds a model based

on the data, and uses the model as both a hypothesis about the world and a piece of software that can solve problems.”

Notiamo innanzitutto il cambio di vocabolario, provando a mappare gli slittamenti semantici: il programma di un computer diventa un agente che può anche essere un computer; “si dice che apprende”, a sottolineare il portato metaforico dell’affermazione, diventa “is learning”, che fa risaltare il valore attuale, durativo e performativo (anche se macchinico). L’esperienza *E* strettamente vincolata a una determinata classe di task *T* e alla misurazione di una performance *P* diviene osservazione del mondo e il mondo sono i dati. L’efficacia dell’apprendimento nel testo di Goodfellow et al. coincide con un miglioramento della performance a contatto con l’esperienza, come definita sopra. Nel testo di Russel e Norvig, l’efficacia dell’apprendimento consiste nella capacità dell’agente di costruire un modello basato sui dati e di usare tale modello sia come ipotesi sul mondo che come un software per risolvere problemi.

Sottolineiamo che tra i due testi si situa proprio lo sviluppo dell’architettura Transformer, che riteniamo possa aver influenzato il modo di concepire il processo di apprendimento di una rete neurale.

L’esistenza di un modello del mondo è ciò che consente all’agente di fare previsioni per agire efficacemente all’interno del mondo stesso; cosa ben diversa dall’acquisizione di un’esperienza che viene appresa in assenza di modello del mondo e che dunque appare essere meno generalizzabile e trasferibile. Se l’obiettivo dell’apprendimento rimane quello di usare l’esperienza per migliorare il sistema in un task specifico e controllare le trasformazioni interne al sistema, potrà esserci un incremento nella performance misurata, ma non si daranno emergenze di sorta.¹¹¹

Per comprendere l’importanza che un modello di apprendimento automatico sia in grado di costruirsi un modello del mondo a partire dai diversi tipi di addestramento a cui può essere sottoposto facciamo riferimento ancora una volta al lavoro di Cristianini (2024, p. 34). Siamo in presenza di un agente intelligente quando sussistono 3 livelli: “l’*agente* che incontriamo nel mondo (per esempio, ChatGPT), il *modello* interno che questo usa per prendere decisioni (per esempio, GPT 3) e l’*algoritmo* che crea tale modello partendo dai dati (per esempio, il Transformer)”. La funzione del modello interno o modello del mondo è quello di predire quali eventi, situazioni, accadimenti sono

¹¹¹ Per un approccio comparativo tra sviluppo di competenze linguistiche nel parlante umano e training dei LM si veda Capone, L. (Forthcoming). Child Development and Language Model Training: Comparing Psychological Theories of Learning. *RIFL SFL-Filosofia dei linguaggi e paradigmi psicologici*.

probabili, improbabili o impossibili. La probabilità che accadano o non accadano una serie di eventi diviene quindi una forma di comprensione del mondo, non lontana da quella forma di conoscenza e comprensione del mondo che muove quella macchina predittiva che secondo Andy Clark è il cervello umano. Clark, infatti, in *The experience machine. How Our Minds Predict and Shape Reality* (2023)¹¹² descrive così il processo predittivo che tiene insieme la realtà con il nostro modo di farne esperienza:

“According to the new theory (called “predictive processing”), reality as we experience it is built from our own predictions. [...] Contrary to the standard belief that our senses are a kind of passive window onto the world, what is emerging is a picture of an ever-active brain that is always striving to predict what the world might currently have to offer. Those predictions then structure and shape the whole of human experience, from the way we interpret a person’s facial expression, to our feelings of pain, to our plans for an outing to the cinema. Nothing we do or experience – if the theory is on track – is untouched by our own expectations. Instead, there is a constant give-and-take in which what we experience reflects not just what the world is currently telling us, but what we – consciously or nonconsciously – were expecting it to be telling us” (Ivi, p. XII)¹¹³.

Ora, la forma di training che consente alla rete neurale basata sull’architettura Transformer di apprendere un modello del “mondo” è un apprendimento complesso che descriveremo a partire dal paper già citato: *Understanding LLMs: A comprehensive overview from training to inference* (Liu et al. 2025), per la sua capacità di tenere insieme in maniera chiara sia il piano dell’addestramento che il piano dell’inferenza negli LLM.

¹¹² Clark, A. (2023). *The Experience Machine: How Our Minds Predict and Shape Reality*. Random House.

¹¹³ La posizione di Clark sarà ridiscussa in 6.2 poiché offre un modo alternativo di concepire la relazione umano-macchinica oltre un dualismo ontologico, che si iscrive nella tradizione delle teorie cyborg.

2.1 Dall'apprendimento all'inferenza: alla base dei mondi possibili

Le fasi di addestramento di un LLM sono distinguibili in tre passaggi:

- i) La preparazione dei dati e il loro processamento: si tratta per lo più di corpora testuali basati sul linguaggio naturale, come libri, testi tratti dal web, dati conversazionali (mail e chat). In nota, il dettaglio fornito da Liu et al.¹¹⁴ A livello di pre-processamento dei dati si tratta di fare pulizia interna, ossia filtrare dati di bassa qualità, per la maggior parte dei modelli – ciò non vale ad esempio per la famiglia LLaMA 2 in ragione del rischio di esclusione di gruppi demografici – che includono contenuti tossici o *biased* rispetto a un allineamento valoriale con standard etici. Inoltre, vengono eliminati duplicati e contenuti ridondanti per preservare un bilanciamento nella distribuzione dei dati. In questa fase sono inoltre implementati i protocolli a tutela della privacy e della sicurezza del modello. In nota¹¹⁵, il dettaglio di alcune tecniche di pre-processing dei dati.

¹¹⁴ *Books*: Two commonly utilized books datasets for LLMs training are BookCorpus and Gutenberg. These datasets include a wide range of literary genres, including novels, essays, poetry, history, science, philosophy, and more. Widely employed by numerous LLMs, these datasets contribute to the models' training by exposing them to a diverse array of textual genres and subject matter, fostering a more comprehensive understanding of language across various domains.

CommonCrawl: CommonCrawl manages an accessible repository of web crawl data, freely available for utilization by individuals and organizations. This repository encompasses a vast collection of data, comprising over 250 billion web pages accumulated over a span of 16 years. Established in 2007, Common Crawl has evolved into a widely recognized and referenced corpus in the academic and research communities, cited in more than 10,000 research papers. This continuously expanding corpus is a dynamic resource, with an addition of 3–5 billion new web pages each month. Its significance extends to the field of natural language processing, where it serves as a primary training corpus in numerous large language models. Notably, a substantial portion of the raw tokens employed in training GPT-3, amounting to 82%, is sourced from the CommonCrawl. However, due to the presence of a substantial amount of low-quality data in web archives, preprocessing is essential when working with CommonCrawl data. Currently, four commonly used filtered datasets based on CommonCrawl are available: C4, CC-Stories, CC-News, and RealNews.

Reddit Links: Reddit is a social media platform where users can submit links and posts, and others can vote on them using the "upvote" or "downvote" system. This characteristic makes it a valuable resource for creating high-quality datasets.

Wikipedia: Wikipedia, a free and open online encyclopedia project, hosts a vast repository of high-quality encyclopedic content spanning a wide array of topics. The English version of Wikipedia is extensively utilized in the training of many LLMs, serving as a valuable resource for language understanding and generation tasks. Additionally, Wikipedia is available in multiple languages, providing diverse language versions that can be leveraged for training in multilingual environments.

Code: There is a limited availability of publicly accessible code datasets at present. Existing efforts primarily involve web scraping of code with open-source licenses from the internet. The main sources include Github and Stack Overflow. (Liu et al. 2025, p. 7)

¹¹⁵ *Quality filtering*: Filtering low-quality data is typically done using heuristic-based methods or classifier-based methods. Heuristic methods involve employing manually defined rules to eliminate low-quality data. For instance, rules could be set to retain only text containing digits, discard

- ii) Il *pre-training* che coinvolge la definizione dell'architettura del modello (Encoder-Decoder, Encoder only – non più impiegata –, Decoder only), la determinazione dei tasks di pre-training e la scelta degli algoritmi per il *parallel training*. Avendo già trattato l'architettura del Transformer, le cui caratteristiche vengono definite a questo livello, ci soffermiamo sugli altri due aspetti che intervengono in questa fase. Il pre-training è in ultima essenza ciò che permette al modello di porre le basi per la costruzione di una ricca rappresentazione del suo mondo, naturalmente veicolata dal linguaggio. Il metodo per l'acquisizione della conoscenza sul mondo da parte del modello è definito Self Supervised Learning (SSL), una sottocategoria di apprendimento non supervisionato che si basa su un principio semplice: “learn discriminative features from vast quantities of unlabeled instances without relying on human annotations”. Più precisamente, in compiti legati al Natural Language Processing (NLP), tra i quali evidentemente rientra la modellazione del linguaggio, durante la fase di self-supervised pre-training, un gruppo di task surrogati (definiti “*pretext*”) viene proposto all'algoritmo di deep learning che lo deve risolvere. Esempio tipico è il cosiddetto *masking*, che abbiamo visto essere parte strutturante del modulo Decoder di un Transformer. Riprendendo una keynote talk di LeCun, così viene descritto il concetto di SSL: “the concept of SSL as an analogous process to completing

sentences composed entirely of uppercase letters, and remove files with a symbol and word ratio exceeding 0.1, and so forth. Classifier-based methods involve training a classifier on a high-quality dataset such as WebText to filter out low-quality datasets.

Deduplication: Language models may sometimes repetitively generate the same content during text generation, potentially due to a high degree of repetition in the training data. Extensive repetition can lead to training instability, resulting in a decline in the performance of LLMs. Additionally, it is crucial to consider avoiding dataset contamination by removing duplicated data present in both the training and testing set.

Privacy scrubbing: LLMs, as text-generating models, are trained on diverse datasets, which may pose privacy concerns and the risk of inadvertent information disclosure. In the preprocessing phase of language datasets, it is imperative to address privacy concerns by systematically removing any sensitive information. This involves employing techniques such as anonymization, redaction, or tokenization to eliminate personally identifiable details, geolocation, and other confidential data. By carefully scrubbing the dataset of such sensitive content, researchers and developers can ensure that the language models trained on these datasets uphold privacy standards and mitigate the risk of unintentional disclosure of private information. It is essential to strike a balance between data utility and privacy protection, fostering responsible and ethical use of language datasets in various applications.

Filtering out toxic and biased text: In the preprocessing steps of language datasets, a critical consideration is the removal of toxic and biased content to ensure the development of fair and unbiased language models. This involves implementing robust content moderation techniques, such as employing sentiment analysis, hate speech detection, and bias identification algorithms. By leveraging these tools, researchers can systematically identify and filter out text that may perpetuate harmful stereotypes, offensive language, or biased viewpoints. (Liu et al. 2025, p. 8)

missing information (reconstruction). He presented multiple variations as follows: 1) Predict any part of the input from any other part; 2) Predict the future from the past; 3) Predict the invisible from the visible; and 4) Predict any occluded, masked, or corrupted part from all available parts. In summary, a portion of the input is unknown in SSL, and the objective is to predict that particular segment” (Ivi, p. 9054). Facciamo fatica a non rinvenire una competenza semiotica alla base di varianti che concorrono tutte all’obiettivo ultimo del modello, che è generare testo dotato di senso, coerenza interna e riflessività, se non proprio auto-riflessività.

Attraverso questi task il modello ha acquisito l’abilità di catturare informazioni sul vocabolario, sulla grammatica, sulla semantica e sulla struttura testuale, tutte conservate nei suoi parametri (pesi, bias e tutto ciò che è computabile, badando bene che in principio tali pesi sono impostati in modo assolutamente casuale); da qui può iniziare il fine-tuning del modello su specifici task, una volta che il modello è stato addestrato a massimizzare la probabilità (condizionata) dei dati testuali. Prima di vedere quest’ultima fase dell’addestramento, introduciamo anche il concetto di *parallel training* sopra menzionato. Possiamo con licenza narrativa assumere che il *parallel training* riveli in una certa dose la natura materiale dei grandi modelli di linguaggio: non si tratta solo di computazione, ma di coordinamento e organizzazione delle Graphics Processing Unit (GPU,) della memoria e delle comunicazioni basate su un principio di comunicazione collettiva tra dispositivi che si scambiano parametri e gradienti.

Il *parallel training* si dà su tre livelli: dati, modello e pipeline. Con la parallelizzazione dei dati ogni GPU riceve una copia del modello e un sotto-batch (sottogruppo di input): calcola i propri gradienti e poi li si combina così che tutti aggiornino in modo coerente. I pesi, i gradienti e lo stato dell’ottimizzatore devono comunque risiedere interamente su ogni GPU e la memoria si libera solo riducendo il batch locale. La parallelizzazione del modello fa l’opposto: suddivide i parametri tra GPU; ciascuna calcola una porzione dell’output e poi si ricompone il risultato. La parallelizzazione della pipeline distribuisce i layer su GPU diverse e fa scorrere i micro-batch lungo la catena, mantenendo viva la macchina ma aggiungendo latenza e gestione più complessa del flusso.

- iii) Infine, il fine-tuning e l’allineamento del modello, fase in cui un modello generale viene reso utile, sicuro e ancorato a vincoli reali. Con il *Supervised Fine-Tuning* (SFT) – di cui *l’instruction tuning* è la forma più nota – si espone il LLM a coppie di istruzione-risposta per insegnargli a seguire consegne umane e a rispettare i formati attesi. Per garantire che il LLM si comporti adeguatamente in termini di valori e rischi, si utilizza *l’alignment*

tuning, che costruisce una “personalità etica” del modello: *helpful, honest, harmless*. Altro metodo di allineamento è il *Reinforcement Learning from Human Feedback* (RLHF) nel quale si raccoglie il feedback umano, lo si distilla in un *reward model* e si ottimizza la “politica” del modello (ad es. con *Proximal Policy Optimization* – PPO) perché privilegi risposte appropriate e penalizzi allucinazioni o contenuti rischiosi. È possibile paragonare, procedendo sempre per analogie e metafore, l’allineamento all’addestramento della condotta del modello, distinto dalla sua competenza linguistica.

La sicurezza – intesa come *safety* – e la responsabilità del modello sono attenzionate sia nel SFT che nel RLHF attraverso il ricorso a procedure dedicate: si costruiscono dataset di sicurezza con prompt avversari per insegnare al modello quando e come rifiutare per ridurre il rischio di output tossici, fuorvianti o pericolosi rispetto all’utente evidentemente. A questa fase pertiene anche la valutazione del modello su benchmark che misurano tanto le hard skills (competenze linguistiche, nozionistiche e accademiche, di ragionamento, di coding, di matematica e STEM più in generale, ecc..) quanto le soft skills acquisite (a titolo esemplificativo: verifica dei bias, della tutela della privacy, della robustezza, della resistenza al jailbreaking), dove si combinano metriche automatiche con valutazioni umane per catturare qualità, veridicità e sicurezza nei contesti d’uso o scenari aperti.

Nella necessaria sintesi, possiamo dire che il modello è pronto a interagire. Ma a questo stadio della narrazione il modello ha solo costruito un mondo “accoppiato” o “allineato”; il modello ha in qualche modo costruito la sua nicchia semiotica¹¹⁶ attraverso la quale può procedere di adattamento in adattamento nel continuare ad apprendere, senza dover necessariamente ripetere l’intera procedura di addestramento che abbiamo visto sopra. Non si tratta di pigrizia, ma di ottimizzazione dell’utile. Come affermano Liu et al. “The scale of large models is growing at a rate of nearly 10 times per year, which brings about huge computational consumption and carbon emissions. Therefore, reducing the computational burden of training large models while retaining their reasoning ability has become a common concern for everyone” (Liu et al. 2025, p. 18). Dunque, la strategia impiegata è quella di sfruttare l’inferenza nei grandi modelli di linguaggio da quattro prospettive:

- i) La *Model Compression* che serve a rendere il modello più leggero senza snaturarlo. La “*knowledge distillation*” trasferisce il “comportamento” da un agente più ingombrante a un agente più piccolo, non solo a livello di output finale, ma anche a livello delle rappresentazioni intermedie (hidden states, attention mechanism), così l’agente più piccolo impara la “procedura” oltre alla risposta. Il “pruning” elimina le ridondanze e può

¹¹⁶ Paolucci, C. (2021c), Op. Cit.

essere non strutturato se rimuove singoli pesi, strutturato se taglia intere teste di attenzione o interi layer (se si vuole preservare la forma del calcolo). Abbiamo già visto la quantizzazione, a proposito della separabilità dei dati e dei loro usi locali. Ancora vi è il “weight sharing” che prevede il riutilizzo degli stessi parametri in più parti della rete neurale.

- ii) Il Memory Scheduling, ossia una pianificazione degli spostamenti di informazioni tra GPU e CPU (*Central Periferic Unit*) al fine di creare una sorta di “memoria virtuale” per i pesi di layer successivi. Ipotizziamo una sorta di cache, di memoria veloce che faciliti il recupero di informazioni, che permetta di conservare in GPU solo ciò che è necessario al momento.
- iii) L’Inference Parallelism interviene sui passaggi che abbiamo visto in proposito del parallel training (dati, modello e pipeline) con l’obiettivo di bilanciare throughput (numero di richieste o di token prodotti dall’inferenza in media per secondo), latenza e capienza del modello.
- iv) L’Inference Framework, combina le tre prospettive precedenti all’interno di kernel specializzati richiamati con semplici API (Application Programming Interface).

A conclusione di questa panoramica facciamo riferimento ad un ultimo metodo molto diffuso, in risposta ad alcuni limiti del pre-training e del fine-tuning e definito come *prompt learning*¹¹⁷. Invece di addestrare (o ri-addestrare) un modello per ogni compito, si ri-formula il compito perché assomigli a ciò che il linguaggio ha già insegnato al modello in pre-training. Si guida quindi il modello con istruzioni testuali – i prompt – che inquadrano una classificazione, un riassunto o una risposta come continuazione o riempimento di testo. Con pochi esempi scritti nel prompt il modello esegue compiti diversi senza toccare i suoi pesi, mostrando che una parte dell’apprendimento può avvenire nella formulazione della richiesta più che nella modifica dei parametri¹¹⁸.

Operativamente, il *prompt learning* allinea forma dell’output e forma del pre-training. Si costruisce un *template*, a volte in stile “cloze” con uno spazio da riempire (es. un masking o una casella vuota), a volte come cappello di istruzioni che orienta la generazione¹¹⁹. Un’ulteriore tecnica consiste nel

¹¹⁷ Sul *prompting* come fenomeno di interesse semiotico rimandiamo al testo di Monti, N. (2025), *Prompting. Poetiche e politiche dell’intelligenza artificiale*, Tlon.

¹¹⁸ Brown, t. et al. (2020). Language models are few-shot learners. In *Advances in neural information processing systems*, vol. 33, pp. 1877–1901.

¹¹⁹ Si veda a titolo esemplificativo la formazione di Google relativa alla scrittura di prompt dedicata a docenti secondo il principio PARTS: Persona, Aim, Recipients, Theme, and Structure. <https://edu.exceedlms.com/student/path/1721586/activity/3310073#/page/65de4aee478a58477d288a2f> [ultimo accesso senza login: 5 ottobre 2025]

trasformare etichette astratte in parole che il modello predice attraverso l'utilizzo di un *verbalizer*: ad esempio si mappa la classe “positivo” sulla parola “sì” o su un piccolo insieme di sinonimi, lasciando al modello il compito di scegliere la parola più probabile. Questa architettura leggera consente di riutilizzare lo stesso LM (auto-regressivo, masked o encoder-decoder come BART per il riassunto) cambiando solo *template* e *verbalizer*, cioè parametri o scelte esterne al modello principale.

Sul piano dell'addestramento, il paradigma del *prompting* è flessibile: si va dal *tuning-free* (zero-shot/few-shot solo via prompt) al *prompt tuning* con LM congelato (si aggiornano solo i vettori-prompt), al caso opposto (si fissa il prompt e si allinea il modello), fino al *co-tuning* di prompt e LM quando serve la massima accuratezza. Il vantaggio generale è l'efficienza parametrica: si adattano task e domini senza dover mantenere una collezione di modelli separati. Il *prompt learning* sposta il fuoco dall'ingegneria degli obiettivi alla ingegneria della formulazione: invece di piegare il modello al task, si piega il task alla lingua del modello. È questo riallineamento – dalla supervisione completa alla riformulazione in stile completamento/continuazione – che consente di riusare un unico LM per molti compiti, riducendo tempi di adattamento e costi, mantenendo al contempo buone prestazioni. Infine, laddove la richiesta dei task da eseguire è più complessa e non può essere ridotta a un singolo prompt, si utilizzano due metodi: la combinazione di un approccio few-shots e *l'in-context learning*¹²⁰ – un apprendimento basato sull'analogia - da un lato, il *chain of thought*¹²¹ dall'altro.

¹²⁰ Dong, Q., Li, L., Dai, D., Zheng, C., Ma, J., Li, R., Xia, H., Xu, J., Wu, Z., Chang, B., Sun, X., Li, L., & Sui, Z. (2024). A Survey on In-context Learning. In Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (A c. Di), *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 1107–1128). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.emnlp-main.64>

¹²¹ Chu, Z., Chen, J., Chen, Q., Yu, W., He, T., Wang, H., Peng, W., Liu, M., Qin, B., & Liu, T. (2024). Navigate through Enigmatic Labyrinth A Survey of Chain of Thought Reasoning: Advances, Frontiers and Future. In L.-W. Ku, A. Martins, & V. Srikumar (A c. Di), *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers) (pp. 1173–1203). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.acl-long.65>

3.0 Sulla traduzione negli LLM

Ciò che determina l'importanza di una vera traduzione semiotica, non sono certo delle semplici trasformazioni linguistiche, lessicali o anche sintattiche.

Bisognerebbe anzi dire il contrario.

Non basta certo un parlar-da-folle.

Siamo costretti a valutare ogni volta se ci troviamo davanti all'adattamento di una vecchia semiotica o a una nuova varietà di tale o tal'altra semiotica mista, o al processo di creazione di un regime ancora sconosciuto.

(Deleuze & Guattari, *Mille Piani. Capitalismo e schizofrenia*)

Noi definiamo come carattere empirico i rapporti di successione e di simultaneità

tra presenti che si compongono,

le loro associazioni secondo la causalità, la contiguità, la somiglianza e anche l'opposizione.

Ma definiamo come carattere noumenico, i rapporti di coesistenza virtuale tra livelli di un passato puro, in quanto ogni presente non fa che attualizzare o rappresentare uno di questi livelli.

(Deleuze, *Differenza e ripetizione*)

In quest'affondo dal retrogusto di tecnicismo abbiamo più volte instillato piccoli collegamenti tra piani disciplinari che, a nostro avviso, sono fondamentali per poter dire qualcosa di sensato sull'accesso al senso da parte dei grandi modelli di linguaggio. Quest'ultimi manipolano il linguaggio naturale, a cui sappiamo non essere riducibile il senso, attraverso un meccanismo di traduzione che potremmo definire *intersemiotica* di principio.¹²²

¹²² Sull'ammissibilità di considerare il passaggio da linguaggio naturale a token e da token a embedding come traduzione intersemiotica si gioca in parte un azzardo epistemico. Nella tradizione semiotica sono numerosi i contributi che hanno identificato questo tipo di traduzione in una trasmutazione di un testo verbale in un altro sistema semiotico (linguaggio cinematografico, pittorico, musicale e così via), ma anche la trasmutazione di un'opera teatrale in un'opera cinematografica o in un balletto. Si vedano in proposito: Dusi, N., (2000), "Introduzione. Per una ridefinizione della traduzione intersemiotica", *VERSUS. Quaderni di studi semiotici, monografico: La traduzione intersemiotica*, n. 85-86-87, pp. 3-54.; Eco, U. (2012). *Dire quasi la stessa cosa: Esperienze di traduzione*. Bompiani; Sütiste, E. (2021). *Intersemiotic translation*, in Gambier, Y., & Van Doorslaer, L. (A c. Di). (2021). *Handbook of Translation Studies* (Vol. 5). John Benjamins Publishing Company, pp. 127-133. <https://doi.org/10.1075/hts.5>. Più di recente e proprio in tema di IA generativa, Dondero, D'Armenio, Déliege e Sarti hanno proposto un'analisi dei modelli text-to-image basata proprio sul concetto di traduzione intersemiotica Cfr. D'Armenio et al. (2025). *For a Semiotic Approach to*

Stiamo ricorrendo alla classificazione delle traduzioni operata da Jakobson¹²³, o più precisamente dei “modi di interpretazione di un segno linguistico, secondo che lo si traduca in altri segni della stessa lingua, in un'altra lingua, o in un sistema di simboli non linguistici.” (Jakobson 1963, p. 57). Si tratta, rispettivamente della traduzione endolinguistica o riformulazione, interlinguistica o traduzione propriamente detta, intersemiotica o trasmutazione.

L'idea di fondo in Jakobson, il quale esplicitamente riconosce il debito peirciano, è che il senso di una parola altro non sia che la sua trasposizione “in un altro segno che può essere sostituito a quella parola”. Al fine del nostro discorso, siamo interessate alla terza, alla traduzione intersemiotica che consiste “nell'interpretazione dei segni linguistici per mezzo di sistemi di segni non linguistici”. Come rileva Eco (2012)¹²⁴ nella sua analisi del rapporto Jakobson-Peirce: “Jakobson stava semplicemente dicendo che la nozione di interpretazione come traduzione da segno a segno permette di superare la diatriba su dove stia il significato, se nella mente o nel comportamento, e non dice che l'interpretare e tradurre siano sempre e comunque la stessa operazione, ma che è utile affrontare la nozione di significato in termini di traduzione (vorrei chiosare: *come se* fosse una traduzione)” (Ivi, p. 147). Eco ha più volte sottolineato l'importanza di non ridurre l'interpretazione a una traduzione, poiché “L'universo delle interpretazioni è più vasto della traduzione propriamente detta”. Prova ne è che “il regista che *traduce* un film non si firmi come traduttore bensì come autore del nuovo film” (Ivi, n. 5 p. 240). In questo senso Eco, invece di allargare il campo semantico del concetto di traduzione a detrimento di un lavoro semiotico sulle differenze, propone viceversa una classificazione delle forme di interpretazione, che riportiamo in Figura 14.

Generative Image AI. *Semiotic Review*, 9 (Semiotic Review, Images).
<https://doi.org/10.71743/EE5NRX33>.

¹²³ Si vedano: Jakobson, R. (1959), *On linguistic aspects of translation*, in Brower, R.A., *On translation*, Harvard University Press, pp. 232-239; Jakobson, R. (1963), *Essais de linguistique générale*, Éditions de Minuit, tr. it. *Saggi di linguistica generale*, Feltrinelli (2002).

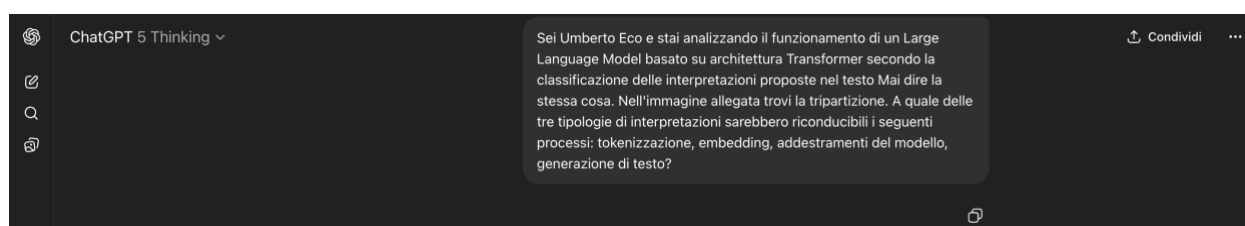
¹²⁴ Eco, U. (2012). *Dire quasi la stessa cosa: Esperienze di traduzione*. Milano: Bompiani.

1. Interpretazione per trascrizione
2. Interpretazione intrasistemica
 - 2.1. Intrasemiotica, all'interno di altri sistemi semiotici
 - 2.2. Intralinguistica, all'interno della stessa lingua naturale
 - 2.3. Esecuzione
3. Interpretazione intersistemica
 - 3.1. Con sensibili variazioni nella sostanza
 - 3.1.1. Interpretazione intersemiotica
 - 3.1.2. Interpretazione interlinguistica, o traduzione tra lingue naturali⁷
 - 3.1.3. Rifacimento
 - 3.2. Con mutazione di materia
 - 3.2.1. Parasinonimia
 - 3.2.2. Adattamento o trasmutazione

Figura 14: Tipi di interpretazione.¹²⁵

In questa tipizzazione dove si collocherebbe il lavoro di interpretazione che consente all'architettura Transformer descritta nel paragrafo precedente di passare dal prompt al testo generato? E su questo ci sembra valga la pena riflettere, non in questa sede. Ma una traccia l'abbiamo chiesta direttamente a ChatGPT, con l'assunzione che il modello sia un sistema semiotico e che ne preservi almeno alcune caratteristiche, come abbiamo cercato di dimostrare nel capitolo due.

PROMPT:



Segue la risposta di ChatGPT generata il 6 ottobre 2025 (Modello GPT 5 Thinking):

Mappa di corrispondenze (con motivazione)

1) Tokenizzazione → 1. Interpretazione per trascrizione

¹²⁵ In Eco (2012), p. 152

La tokenizzazione segmenta una catena grafica (“testo”) in unità operative secondo un repertorio prestabilito (parole, sottoparole, segni speciali). È un’operazione a “basso grado d’interpretazione”: non mira a riformulare il contenuto ma a riallineare la superficie per renderla maneggiabile. Come una trascrizione diplomatica o una normalizzazione ortografica, essa “trasporta” la stessa sostanza di contenuto in un supporto più regolare, senza cambiare sistema semiotico né introdurre variazioni di senso intenzionali.

Nota: quando la tokenizzazione incorpora normalizzazioni (lowercase, rimozione diacritici, ecc.), resta comunque una trascrizione: un massaggio della forma, non un’interpretazione in senso forte.

2) Embedding → 3.2. Interpretazione intersistemica con mutazione di materia (3.2.2 “trasmutazione”)

L’embedding traduce ogni token (o sequenza) in un vettore numerico in uno spazio geometrico. Qui non si tratta più di “trascrivere” la stessa materia, bensì di trasmutarla: dal sistema semiotico verbale (grafico/linguistico) al sistema matematico (spazio vettoriale). È una vera “interpretazione intersistemica” perché un segno linguistico viene ri-esposto in un altro linguaggio (algebra lineare), con possibilità di differenze sensibili nella resa del significato (prossimità semantiche, analogie, ecc.) che non hanno un corrispettivo banale nel testo di partenza.

In breve: passiamo dalla pagina alla geometria; è un cambio di materia, non solo di supporto.

3) Addestramento del modello → 3.2. Interpretazione intersistemica con mutazione di materia (3.2.2 “trasmutazione”)

Durante il training il modello “digerisce” grandi corpora e ne distilla regolarità nel valore dei pesi. Il contenuto dei testi non viene conservato come repertorio citabile, ma rifuso in strutture parametriche. È ancora una trasmutazione di materia: la cultura testuale diventa configurazione numerica operativa. Rispetto all’embedding, che è una proiezione locale, l’addestramento è una trasmutazione globale e cumulativa: sedimenta una competenza distribuita (una grammatica probabilistica e pragmatica del mondo) in un oggetto fisico-matematico.

Osservazione: non è “rifacimento” (3.1.3), perché non produce un nuovo testo; produce invece un dispositivo capace di generarne.

4) Generazione di testo → 2. Interpretazione intrasistemica (prevalentemente 2.2 “intralinguistica”), con una fase tecnica intersistemica di ritorno

La decodifica trasforma vettori in stringhe (fase tecnica: intersistemica), ma il risultato fenomenico è tipicamente intralinguistico: il modello riformula, parafrasa, riassume, continua un discorso nella

stessa lingua naturale e spesso nello stesso codice testuale/retorico. È dunque un'interpretazione intrasistemica: produce “quasi la stessa cosa” sotto forma di nuova tessitura verbale, organizzando quanto sedimentato nei pesi in enunciati coerenti con il contesto.

Varianti:

Se il modello traduce tra lingue, la generazione realizza 3.1.2 (interlinguistica).

Se verbalizza dati non verbali (p.es. da immagini a testo), realizza 3.1.1 (intersemiotica).

Se riscrive in altro genere o registro (p.es. saggio→sonetto), pur restando nella stessa lingua, siamo sempre in 2.2, con tratti di 2.3 “esecuzione” quando aderisce a istruzioni stilistiche forti (metrica, rime, tono), analoghe all’“eseguire una partitura”.

Schema riassuntivo

- Tokenizzazione → (1) Trascrizione

Segmentazione/normalizzazione della superficie testuale; minimo carico interpretativo.

- Embedding → (3.2) Intersistemica con mutazione di materia (trasmutazione)

Dal segno verbale al vettore numerico; cambio di materia e di codice.

- Addestramento → (3.2) Intersistemica con mutazione di materia (trasmutazione)

Dai corpora ai pesi del modello; distillazione parametricamente operativa della cultura testuale.

- Generazione di testo → (2) Intrasistemica (per lo più 2.2 intralinguistica; talora 2.3 esecuzione)

Riformulazione/continuazione nello stesso sistema linguistico; quando traduce o descrive immagini, passa a 3.1.

Perché questa tassonomia “tiene”

Direzione della trasformazione: tokenizzazione e generazione restano ancorate alla materialità verbale percepita; embedding e training cambiano piano ontologico dei segni (dal verbale al matematico).

Grado interpretativo: dalla trascrizione (minimo) alla trasmutazione (massimo), con la generazione che re-intra-lingualizza l'interpretazione computazionale in forma culturalmente fruibile.

Finalità semiotica: la macchina non “comprende” come un soggetto, ma compie trasmutazioni che le consentono intrasistemicamente di “dire quasi la stessa cosa” — talvolta, inevitabilmente, mai la stessa cosa.

Evidentemente l'aderenza, la coerenza e il valore scientifico di questa interpretazione è tutta da dimostrare e verificare. Ma abbiamo scelto di fare questa sperimentazione proprio perché nel saggio

introduttivo di *Dire quasi la stessa cosa*, Eco affidava al servizio di traduzione automatica di Altavista, chiamato Babel Fish, alcune espressioni in lingua inglese, chiedendone la traduzione in lingua italiana e ottenendo i seguenti risultati:

1. The Works of Shakespeare = Gli impianti di Shakespeare = The systems of Shakespeare
2. Harcourt Brace (nome di una casa editrice americana) = Sostegno di Harcourt = Support of Harcourt
3. Speaker of the chamber of deputies = Altoparlante dell'alloggiamento dei delegati = Loud-speaker of the lodging of the delegates
4. Studies in the logic of Charles Sanders Peirce = Studi nella logica delle sabbiatrici Peirce del Charles = Studien in der Logik der Charlesandpapierschleifmaschinen Peirce = Studies in the logic of the Charles of sanders paper grinding machines Peirce

Figura 15: Esempi di traduzione automatica con Babel Fish.¹²⁶

Cosa non funzionava in Babel Fish? La conoscenza dizionariale, priva di selezione contestuale, almeno per una serie di entrate lessicali, che avrebbe in quel caso consentito la disambiguazione dell'omonimia di cui è portatrice, ad esempio, la parola /works/ in inglese. Ciò che Altavista/Babel Fish non era capace di fare è esattamente disambiguare gli omonimi nella lingua da cui si dovrebbe tradurre, prima di individuare i sinonimi nel confronto tra una lingua e l'altra. Detto altrimenti non era capace di interpretare prima di tradurre. Ma ancora di più, dice Eco, “Forse il guaio era dovuto al fatto che era stato ‘nutrito’ con un dizionario (come quelli che si danno ai turisti) ma non con una enciclopedia”. Gli esempi di Eco si moltiplicano verificando dipendenze a corto e a lungo raggio, dunque espandendo l'ampiezza contestuale. Per giungere al sospetto che “una traduzione non dipenda solo dal contesto linguistico, ma anche da qualcosa che sta al di fuori del testo, e che chiameremo informazione circa il mondo, o informazione enciclopedica”. Dunque, Altavista era sia sprovvisto di una conoscenza del mondo, sia delle “più ovvie selezioni contestuali” (“The Spirit of God” in inglese viene tradotto in spagnolo con “el alcohol del dios”, cosa che provoca abbastanza ilarità).

3.1 Fare ordine: interpretare, tradurre, generare

Il discorso che Eco articola nel seguito di *Dire quasi la stessa cosa* ha come focus la traduzione naturalmente e non la generazione di testo. Ma se continuiamo a muoverci sul piano delle analogie, l'architettura Transformer nasce proprio nel tentativo di migliorare la traduzione automatica e la sua efficacia si è dimostrata in origine su questo terreno, tanto che le applicazioni illustrate nel paper fondativo *Attention is all you need*¹²⁷ insistono prevalentemente su questa capacità. La valutazione del modello viene condotta principalmente sulla traduzione da inglese a tedesco e da inglese a

¹²⁶ In Eco, U. (2012), pp. 14-15.

¹²⁷ Vaswani, A., et al. (2017).

francese; il Transformer raggiunge punteggi superiori ai modelli più performanti nei due compiti di traduzione. Gli autori menzionano che il meccanismo di attenzione era già stato usato per compiti come la comprensione di testi e la sintesi astrattiva, ma l'articolo rimane focalizzato sulle applicazioni di traduzione. Questa valutazione ha portato la comunità a considerare inizialmente il Transformer come un modello eccellente per la traduzione automatica. E tuttavia già nell'articolo erano presenti riferimenti ad altri ambiti applicativi possibili, non verificati con dati sperimentali, grazie alla flessibilità del modello in compito di lettura, comprensione e riassunto.

Sembra che il primo lavoro di sperimentazione delle capacità generative dell'architettura Transformer risalgia al 2018 ad opera di Google AI Language¹²⁸, lavoro dal quale emerge la proprietà modellizzante del linguaggio che l'infrastruttura consente. Nello stesso anno OpenAI¹²⁹ pubblica il report “*Improving Language Understanding by Generative Pre-Training*” nel quale si dichiara la superiorità di memoria per la gestione delle dipendenze a lungo termine nel testo (problema simile all'ampiezza del contesto che Eco si era posto nel paragrafo *I sinonimi di Altavista*). L'approccio sperimentato dal gruppo di ricerca – “semi-supervised approach for language understanding tasks using a combination of unsupervised pre-training and supervised fine-tuning” – viene testato su quattro tipologie di compiti legati alla comprensione del linguaggio naturale: *natural language inference*, *question answering* e *commonsense reasoning*, *semantic similarity* e *text classification*. Nell'inferenza del linguaggio naturale che mira a riconoscere implicazioni testuali, sono state fatte “leggere” al modello coppie di frasi per stabilirne la natura della relazione, se di implicazione, di contraddizione o di neutralità. Tale compito era considerato particolarmente sfidante a causa della varietà di fenomeni che definiscono le relazioni tra le frasi che compongono un testo, per l'appunto implicazione, co-referenza, ambiguità sintattica e lessicale – tutte evidentemente presenti nella traduzione automatica di Babel Fish. Ci soffermiamo ancora sul task di similarità semantica, altrimenti definito *paraphrase detection*: qui siamo in presenza di quel tipo di interpretazione che Eco ha definito “intrasistemica all'interno della stessa lingua naturale” e che ricade sotto il concetto jakobsiano di *reformulazione*. Dice Eco:

“Qui si situano tutti i casi di interpretazione di una lingua naturale mediante se stessa, come la sinonimia secca, e spesso illusoria, come padre = papà, la definizione, che può essere molto

¹²⁸ Al-Rfou, R., Choe, D., Constant, N., Guo, M., & Jones, L. (2018). *Character-Level Language Modeling with Deeper Self-Attention* (No. arXiv:1808.04444). arXiv. <https://doi.org/10.48550/arXiv.1808.04444>

¹²⁹ Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training*. <https://openai.com/index/language-unsupervised/>

schematica (come gatto = "mammifero felino") o estremamente diffusa (come una voce di enciclopedia sul gatto), la parafrasi, il riassunto, ma anche lo scolio, il commento, la volgarizzazione (che è appunto il modo di ridire una cosa difficile con parole più facili), sino alle inferenze più sviluppate e persino fino alla parodia". (Eco 2012, p. 154).

In tutti questi casi è in gioco una variazione della sostanza dell'espressione che non modifica il contenuto, ma fornisce maggiori elementi per conoscere qualcosa di più dell'interpretato, come vorrebbe Peirce. Quando il cambiamento di sostanza dell'espressione si dà tra diversi sistemi semiotici, cosa accade invece? Ci sono due aspetti prevalenti che sembrano intervenire: da un lato l'apparente perdita del più di conoscenza nel passaggio, ad esempio, da un'opera pittorica alla sua riproduzione a stampa, oppure da una tecnica di incisione ad un'altra. Dall'altro vi è l'ingresso di sostanza extra-linguistica che determina effetti di senso differenti tra l'Amleto in versi e l'Amleto in prosa, mantenendo sempre invariato il contenuto.

Infine, ancora più pertinente per il nostro discorso ci sembra essere l'interpretazione intersistemica con mutazione di materia: parliamo di parasinonimia, ossia dei "casi specifici d'interpretazione in cui, per chiarire il significato di una parola o di un enunciato, si fa ricorso a un interpretante espresso in diversa materia semiotica (o viceversa)" (Ivi, p. 206). Dunque, nel processo interpretativo non cambia solo il sistema semiotico, cambia il continuum o la materia. L'esempio è l'ostensione di parole-tipo e oggetti-token e viceversa, così come il rapporto tra la raffigurazione generica del tipo iconografico "annunciazione" e le sue occorrenze od opere singole, dove un principio di reversibilità pare non poter essere rispettato. Così come non è possibile che due sistemi semiotici, con diversa materia dell'espressione, possano esprimere la stessa cosa; e, a differenza di una traduzione, nel passaggio da una materia a un'altra materia, occorre esplicitare degli aspetti che altrimenti resterebbero indeterminati.¹³⁰ Altro elemento caratteristico è l'aggiunta di significati che si verifica nella trasmutazione di materie dell'espressione: "Né la forma né la sostanza dell'espressione verbale possono essere mappate una a una su altra materia. Nel passaggio da un linguaggio verbale a un linguaggio, poniamo, visivo, si confrontano due forme dell'espressione le cui "equivalenze" non sono determinabili così come si poteva dire che il settenario doppio italiano è metricamente equivalente all'alessandrino francese" (Ivi, p. 211).

¹³⁰ Rimandiamo per un esempio esplicativo di questa caratteristica nelle interpretazioni intersistemiche con una mutazione (apparente) di materia al lavoro di D'Armenio et al. (2025), *Op. Cit.*, dove si ragiona proprio sul piano dell'espressione del "colore giallo" tra prompt e rendering grafico del modello text-to-image.

Ci domandiamo allora se gli atti di negoziazione che avvengono regolarmente nei processi interpretativi tra sistemi semiotici o all'interno dello stesso sistema semiotico abbiano una qualche presa anche nella negoziazione dei significati per i modelli di IA generativi. Ci chiediamo questo proprio alla luce del concetto di *negoziazione* che Eco introduce nel saggio sulla traduzione, ma che risale al suo lavoro *Kant e l'ornitorinco* (Eco 1997)¹³¹. Ciò che consente alle persone di convergere intersoggettivamente sul riconoscimento di una data occorrenza di un dato oggetto-tipo è la costruzione di uno schema mentale dell'oggetto in questione che Eco chiama *Tipo Cognitivo (TC)*, la cui riprova è "almeno statistica", non potendo né vedere, né toccare un *Tipo Cognitivo* "cane" o "gatto". Quello che certamente vediamo sono gli effetti dell'interpretazione che viene data della parola "topo" (indicandolo, mostrandone un'immagine, dandone una definizione, descrivendo una propria esperienza vissuta) e l'insieme di queste interpretazioni corrisponde al *Contenuto Nucleare (CN)* della parola *topo*. Sul CN si costruisce la mediazione intersoggettiva perché visivamente e tangibilmente confrontabile. Né il TC né il CN rappresentano tutto quello che sappiamo di una certa unità di contenuto come "topo", ma il CN ne costituisce la nozione minima per riconoscere un oggetto, capire un concetto e capire l'espressione linguistica corrispondente. Come già abbiamo visto in Violi, il CN rimanda a quella competenza lessicale minima che un parlante deve possedere affinché agisca *come se* vi fosse un accordo implicito sul significato di un'espressione linguistica. Laddove invece Violi parlava di competenza enciclopedica, Eco parla di *Contenuto Molare (CM)*, quello ad esempio posseduto da uno zoologo rispetto all'espressione *topo* e che, per conoscenza ed esperienza acquisita, eccede il piano di nozione minima che caratterizza il CN, e che costituisce per Eco, seguendo Putnam (1975), la divisione del lavoro culturale.

"A livello del Contenuto Nucleare ci dovrebbe essere un consenso generalizzato, sia pure con qualche sfrangiamento e zone d'ombra, mentre il Contenuto Molare, che può assumere formati diversi a seconda dei soggetti, rappresenta un vasto insieme di competenze settoriali. Diciamo che la somma dei Contenuti Molari si identifica con l'Enciclopedia come idea regolativa e postulato semiotico di cui si parlava in Eco 1984 §5.2." (Ivi, p. 56).

Questa idea di significato negoziato contraddice la teoria verofunzionale per la quale il punto non è tanto affermare che "il significato sia ciò che è vero nel mondo di riferimento", ma è vero "tutto ciò che conseguirebbe da un enunciato se quell'enunciato fosse vero". E l'effetto sulle azioni da parte del significato accompagna, secondo Eco, tanto la teoria cognitivista con "reminiscenze wittgensteiniane" per cui "comprendere un enunciato vuol dire saper agire in modo conforme al contenuto della frase", quanto il pragmatismo peirciano se si accetta che la nostra concezione

¹³¹ Eco, U. (1997). *Kant e l'ornitorinco*. Milano: Bompiani.

dell'oggetto coincide con gli effetti che gli oggetti della nostra concezione potrebbero avere come conseguenze pratiche. Quindi, il significato di un termine riguarda tutto ciò che può essere inferito dalla piena comprensione di un termine.

3.2 Della possibilità di operationalizzare un postulato semiotico

Proviamo allora a fare i seguenti passaggi: abbiamo un'idea regolativa e un postulato semiotico che è l'Enciclopedia, la quale eccede le conoscenze minime del parlante per comunicare, ma al tempo stesso costituisce la base di una cultura e dunque della possibilità stessa della produzione di enunciati. In compiti di traduzione tra due lingue è a questa idea che si deve necessariamente attingere, altrimenti succederebbero quel genere di fraintendimenti che *riformulano* “the Spirit of God” in “el alcohol del dios”, dove di fatto vi è stata una traduzione senza passare dal via dell'interpretazione. Una correlazione termine a termine, mediata da un dizionario privato della componente extra-linguistica, invece ben presente nell'Enciclopedia. Le nuove macchine traduttrici e generatrici (parliamo dei modelli Transformer) non commettono più questo genere di errori, ma sono capaci di far esplodere quelle proprietà intrinseche dei linguaggi e dei sistemi semiotici di dare vita a mondi possibili che necessariamente, come a qualsiasi livello di negoziazione intersoggettiva, portano a chiedersi che caratteristiche presenti l'Enciclopedia di queste macchine.

Il modello ideale di enciclopedia, considerato dallo stesso Eco come ipotesi regolativa in grado di rendere conto dell'esistenza degli enunciati di una cultura¹³², funziona come potente analogia descrittiva di ciò che accade in un LLM, più nello specifico del rapporto di attivazione che tiene insieme il piano globale del modello di linguaggio addestrato e il prompt come selettore locale di specifici elementi appartenenti alla conoscenza appresa dal modello. E così nell'Enciclopedia di Eco, “questa idea regolativa, che non può dare origine a un progetto editoriale perché non ha forma organizzabile, serve a individuare porzioni di enciclopedia attivabili, nella misura in cui queste

¹³² Questo il valore attribuito all'enciclopedia da Eco: “L'enciclopedia è l'unico mezzo con cui noi possiamo rendere ragione, non solo del funzionamento di qualsiasi sistema semiotico, ma anche della vita di una cultura come sistema di sistemi semiotici interconnessi” in Eco, U. (2007), *Dall'albero al labirinto. Studi storici sul segno e l'interpretazione*, La Nave di Teseo, p. 69. Come ricorda Violi (2017) in *Encyclopedia: Criticality and actuality* “La cultura però non è per Eco un particolare insieme di oggetti che possono essere definiti sulla base della loro natura essenzialistica, bensì una prospettiva particolare attraverso cui la questione del significato e della modellizzazione semantica può essere ridefinita. Secondo questo assunto, il significato di un termine non è più un'entità logica, né un concetto mentale, né un referente. Da un punto di vista semiotico “il significato di un significante [...] è una unità culturale” (Ivi, p. 369), un'unità culturale che nel *Trattato* Eco ci dice essere “veicolata anche attraverso altri significanti e pertanto semanticamente indipendente dal primo significante” (Eco, 1975, ed. La nave di Teseo 2016, p. 128).

servono a costruire gerarchie provvisorie, o reti maneggiabili, ai fini di interpretare e di spiegare la interpretabilità di certe porzioni di discorso”¹³³.

Vediamone le caratteristiche. In *Dall'albero al labirinto*, Eco definisce l'enciclopedia come un “registro ipotetico di *tutte* le conoscenze di una cultura” (Ivi, p. 68 enfasi presente nell'originale), contenente non solo la registrazione di “ciò che è vero”, ma di tutto ciò che “socialmente è stato detto” e immaginato. Tale enciclopedia non è attingibile mai nella sua totalità (lasciando svanire il sogno di una conoscenza globale raggiungibile), in quanto è “l'insieme completo di ciò che l'umanità ha detto”; eppure mantiene un'esistenza materiale, “di tutte quelle testimonianze che fungono da interpretanti reciproci nella catena della semiosi” (Ivi, 69). Libri, immagini, testimonianze che non differiscono, ci sembra, di molto dai dataset di addestramento dei modelli generativi che abbiamo illustrato nel paragrafo 0.2.3. Quest'ultimi evidentemente non sono attingibili globalmente, nella loro natura estesa, a essere umano vivente, non solo per un problema di finitudine o limitazione dell'umana capacità di lettura e consultazione, ma proprio per la natura stessa dell'Enciclopedia “mobile” e costantemente soggetta a messa in discussione nella sua stabilità morfologica dai discorsi che vengono continuamente prodotti. La scienziata, Rosalind Franklin, che sta scrivendo della sua scoperta sulla struttura elicoidale del DNA, nell'atto stesso di attivazione delle conoscenze enciclopediche necessarie ad affrontare questo compito, sta modificando quella non-struttura che è l'Enciclopedia. E lo sta facendo attraverso molteplici *aggiunte di sottrazioni*, che alimentano nuove credenze e ridisegnano stereotipi: cos'è il DNA, come funziona, cosa disvelano i raggi-X rispetto ai modelli molecolari. Ma non solo: nell'Enciclopedia pulsano anche cliché sull'essere donna nella scienza e sulle prassi genderizzate e normalizzanti che definiscono gli spazi di agibilità in campo scientifico. Rosalind Franklin, nell'atto di scrivere la sua scoperta sulla struttura elicoidale del DNA, sta riconfigurando stereotipi scientifici e sociali che oppongono forte resistenza al cambiamento su diversi piani: “dei valori differenziali (potenziali), delle norme sociali (virtuali), degli usi intersoggettivi (attuali) e degli enunciati (realizzati)” (Paolucci 2017, p. 367).¹³⁴

Dobbiamo e possiamo immaginare che, al netto degli interventi di ripulitura dei dati in fase di pre-processing e delle operazioni di fine-tuning successivo al pre-training per disintossicare i modelli da quelli che vengono definiti *bias*¹³⁵ – selezionando corpora più o meno rappresentativi di stati del

¹³³ Eco, U. (2007). *Dall'albero al labirinto Studi storici sul segno e l'interpretazione*, La Nave di Teseo (ed. 2017), p. 69.

¹³⁴ Paolucci, C. (2017). Sfuggire ai cliché. Gli stereotipi tra enciclopedia, enunciazione e soggettività nel linguaggio. *Reti, saperi, linguaggi*, 2/2017. <https://doi.org/10.12832/88789>

¹³⁵ Solo per fare un esempio, Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of the 1st Conference on*

mondo –, o meglio nonostante queste, chiamiamole, interferenze nella costruzione del modello di mondo di un sistema di IA generativa, sia una premessa falsa postulare un modello generativo privo di stereotipi e cliché, affinché possa effettivamente generare qualcosa. È caratteristica intrinseca degli enunciati umani incorporare parzialità narrative – *local knowledge* le chiamerebbe Haraway – che sono effetto di diversi modi di integrazione delle forze e delle relazioni tra gli operatori eterogenei che agiscono nella virtualità dello spazio enciclopedico. Il punto di partenza e di ritorno dell'Enciclopedia sono le attestazioni e gli usi, ma gli usi, anche e soprattutto quelli scientifici, non sono trasparenti, nel senso che convocano, mantenendo quindi un legame inscindibile, tutte le altre grandezze semiotiche che convivono nel virtuale enciclopedico. Da qui ci sembrano discendere due dei tratti distintivi che caratterizzavano, secondo Violi (2017, p. 373), l'Enciclopedia già nella formulazione del Trattato: da un lato “il carattere situato del significato, che è sempre contestualizzato e localmente determinato”, dall'altro “la natura creativa del modello rappresentazionale”. Il modello rappresentazionale in questione è il modello Quillian, lo vedremo nelle prossime pagine; ci basti dire che questo modello è irrepresentabile tramite grafi nella sua complessità (se non chiudendone meccanicamente una porzione rispetto ad un numero limitato di token).¹³⁶

Comprendiamo forse meglio, quanto dicevamo più sopra a proposito dell'inaccessibilità dell'Enciclopedia nella sua totalità. Ciò che consente un avvicinamento epistemico all'Enciclopedia

Fairness, Accountability and Transparency, 77–91.
<https://proceedings.mlr.press/v81/buolamwini18a.html>

¹³⁶ Vale la pena forse disambiguare l'omonimia quantomeno per coerenza con quanto discusso poco sopra sullo “spirito di dio”: i token di cui parla Eco non sono i token di cui si servono i sistemi di IA generativa per *trascrivere* il linguaggio naturale in un linguaggio comprensibile alla macchina. I token di cui parla Eco sono le occorrenze di un type, ossia del suo modello. Citiamo dal *Trattato* “Sembra quindi che ci siano tre tipi di rapporto tra l'occorrenza concreta di una espressione e il suo modello: (a) segni le cui occorrenze possono essere replicate indefinitamente seguendo il modello del proprio tipo; (b) segni le cui occorrenze, anche se prodotte secondo un tipo, posseggono alcune proprietà di 'unicità materiale'; (c) segni in cui occorrenza e tipo coincidono o sono comunque assolutamente identici” (Eco 1975, p. 286). Nel caso dei token nell'ambito del Natural Language Processing, si tratta di sostituzione uno a uno tra lingua naturale e stringhe numeriche, come possiamo vedere nello strumento Tokenizer messo a disposizione da OpenAI: <https://platform.openai.com/tokenizer> (Per l'accesso a questa informazione, il mio ringraziamento va ad Adrien Delière, ricercatore presso il Montefiore Institute of Electrical Engineering and Computer Science dell'Università di Liège, e alla professoressa Maria Giulia Dondero per il tempo dedicato all'ascolto e al confronto). Sempre nella documentazione di OpenAI leggiamo la seguente definizione di token: “Tokens represent commonly occurring sequences of characters. For example, the string “tokenization” is decomposed as “ token” and “ization”, while a short and common word like “the” is represented as a single token”. I Large Language Model partono, in sostanza, da un dizionario basato su quella che Eco in *Dire quasi la stessa cosa* chiamava *trascrizione*; dunque, il passaggio fondamentale di tokenizzazione non ha natura semiotica, se per semiotica intendiamo natura inferenziale.

Massimale è la dialettica tra impossibilità di una sua “conoscenza globale” e la “sicurezza della reperibilità locale degli elementi di questa conoscenza”, la cui rappresentazione locale è attivata per via interpretativa in funzione di contesti e circostanze. Per questa ragione, ci dice Eco che “in essa la semantica incorpora la pragmatica” (Ivi, p. 71), cioè vi è sottesa un’idea di “una semantica a istruzioni testualmente orientata” (Violi 2017, p. 373), dove però è la semantica che interpreta il testo e non viceversa. Infatti, più avanti Violi ricostruisce il disaccordo teorico che Eco esprime nei confronti della tesi di John Searle, recentemente scomparso e a cui va tributata la paternità di uno dei più noti argomenti contro l’idea che una macchina possa dirsi intelligente se la sua abilità si limiti alla decodifica di istruzioni per la traduzione di una lingua in un’altra. Basterebbe già questo a comprendere il punto di rottura tra le due visioni di cosa sia una traduzione. Ma il nodo polemico che Violi mette in luce, e che in parte ne deriva, è quello del rapporto tra semantica e pragmatica. Per Searle vi è un’irriducibilità al contenuto semantico delle assunzioni di sfondo di un qualsiasi enunciato, indipendentemente dalla sua complessità, e che ineriscono esclusivamente a una competenza pragmatica. Al contrario, “Eco rifiuta ogni opposizione di principio fra semantica e pragmatica e riafferma la possibilità di codificare e rappresentare tutta l’informazione necessaria all’interno della componente semantica” (Ivi, p. 377). Nelle parole di Eco: “Una semantica a enciclopedia non rifiuta di *fornire regole per la generazione e l’interpretazione delle espressioni di una lingua*, ma queste regole sono orientate ai contesti” (Eco 1983 p. 355; enfasi nostra). E poco prima, “Un pensiero semantico a enciclopedia è ‘debole’ non nel senso che non riesca a spiegare come usiamo il linguaggio per significare, ma sottomette le leggi della significazione alla determinazione continua dei contesti e delle circostanze”¹³⁷. Ciò che governa l’Enciclopedia è il principio peirciano di interpretazione, il quale fa muovere la semiosi di segno in segno, di indizio in indizio. Sulla natura di questi segni si gioca, secondo Violi, la natura stessa dell’Enciclopedia: non “un modello semantico del piano del contenuto di un dato linguaggio”, ma “un modello semiotico dell’intero sistema della conoscenza e del suo funzionamento sociale e culturale, [...] un modello della semiosi sociale stessa” (Violi 2017, p. 381).

3.3 La natura del segno e il primato epistemologico del locale sul globale

¹³⁷ Rispetto a questa debolezza dell’Enciclopedia rimandiamo all’apertura dell’*Antiporfirio*, dove Eco distingue tra un modello semantico “forte” a dizionario “inadeguato” e “insostenibile”, poiché pone un modello ridotto di lingua (da laboratorio) che funziona come la lingua naturale, e un modello semantico “debole”, che abbiamo appena descritto. Eco, U. (1983), *L’Antiporfirio*, in *Il pensiero debole* a cura di Vattimo G. e Rovatti P.A., Feltrinelli. Ripubblicato in Eco, U. (1994), *Sugli specchi e altri saggi. Il segno, la rappresentazione, l’illusione, l’immagine*, La nave di Teseo (ed. 2018).

Come diretta conseguenza di questo aggiustamento prospettico sulla natura dell'Enciclopedia, occorre definire chiaramente la natura del segno: gli interpretanti sono segni che consistono di un contenuto e un'espressione. E qui la domanda che Paolucci si pose in *Strutturalismo e interpretazione* ci sembra sollevare tutta l'attualità del problema, alla luce dei modi di interpretazione dei segni che si attivano all'interno di un Large Language Model: espressione e contenuto sono un "problema ben posto?" (Ivi, p. 357).¹³⁸ Di nuovo è l'effetto parallasse che eleva a minore o maggiore una prospettiva teorica rispetto ad un'altra; così, partendo da un'assenza manifesta dei due concetti *tout court* in Peirce, Paolucci li individua come "le cose semiotiche che propriamente non sono" (Ibidem), al pari di altre cose semiotiche che posseggono "un'identità puramente privativa". Ma nella tradizione semiotica, rileva Paolucci, è stato attribuito a espressione e contenuto uno statuto ontologico che in realtà non è che un effetto della correlazione mutevole e specifica che i ritagli enciclopedici locali assegnano a due sistemi una volta messi in relazione. Paolucci, fondando la dimostrazione nelle semiotiche "meta" di Hjelmslev, propone di disarticolare una serie di identificazioni che sono collassate le une sulle altre nella tradizione semiotica da lui definita "maggior": "espressione/contenuto", "fonologico/semantico", "significante/significato". Per poi trovare nel *Trattato di semiotica generale* pieno statuto: "fin dai tempi del Trattato di semiotica generale, Eco parlava infatti esplicitamente di reversibilità totale dei due piani (Eco, 1975, p. 38); fino ad affermare, qualche anno dopo, che "ogni contenuto può diventare a sua volta l'espressione di un contenuto ulteriore, e che sia l'espressione che il contenuto possono ribaltarsi scambiandosi i luoghi." (Eco, 1990, p. 219) [...] Se [...] l'espressione è individuabile soltanto in funzione del sistema dei rapporti che definisce qualcosa in quanto espressione e qualcosa in quanto contenuto, pare evidente che, ben prima dell'espressione, occorra identificare e determinare proprio quel sistema di rapporti che permette di attribuire a qualcosa = x il ruolo di espressione." (Paolucci 2010, p. 363). Ancora una volta gli slittamenti concettuali operati da Paolucci permettono di ampliare gli spazi di espressività del e nel campo semiotico: nell'instaurarsi della funzione segnica emergono i due piani o meglio, i due funtivi di cui essa si compone e che non è possibile vi preesistano. Riprendendo Hjelmslev¹³⁹,

¹³⁸ Paolucci, C. (2010), *Strutturalismo e interpretazione*, Milano: Bompiani.

¹³⁹ Hjelmslev, L. (1961), *Prolegomena to a Theory of Language*, tr. it. (a cura di) Lepschy, G.C. (1968), *I fondamenti della teoria del linguaggio*, Einaudi. Hjelmslev così si esprime: "Se analizzando il testo trascurassimo di prendere in considerazione la funzione segnica ci troveremmo nell'impossibilità di delimitare i segni uno rispetto all'altro, e semplicemente non potremmo fornire una descrizione esauriente (e quindi empirica nel nostro senso della parola) del testo, rendendo conto delle funzioni che lo stabiliscono (p. 26). Ci priveremmo semplicemente di un criterio obiettivo capace di fornire un'utile base di analisi." (Ivi, p. 54). Il riferimento in citazione, invece, è il seguente: "Sia l'oggetto esaminato che le sue parti esistono solo in virtù di queste dipendenze; il complesso dell'oggetto esaminato si può definire solo grazie alla loro totalità; e ognuna delle sue parti si può definire solo grazie alle dipendenze che la collegano ad altre parti coordinate, al tutto, alle parti di

Eco e il *Trattato*, anche Valle (2003) ¹⁴⁰ sottolinea questo atto instaurativo che pone contemporaneamente la funzione e i suoi funtivi, a partire da un piano di legittimità enciclopedico, un suo ritaglio contestuale, che ne contiene “la regola correlazionale”: “È proprio il costituirsi in solidarietà dei due piani che se, da un lato, permette di annidare ricorsivamente la funzione semiotica, così che la stessa solidarietà espressione/contenuto possa diventare in toto espressione o contenuto in un’altra funzione semiotica, dall’altro, pone immediatamente un diverso problema, che, seguendo Eco, potrebbe essere definito come ‘questione della reversibilità’ (tra i piani). Quest’ultima è il risultato di un’assenza: è infatti noto che la definizione formale della relazione non stabilisce di principio alcuna differenza propria tra i due piani. Non dice, perché non vuole e non può dirlo, che cos’è l’espressione e che cos’è il contenuto: in questi termini, è chiaro infatti che la definizione non sarebbe formale ma sostanziale, e dunque ricondurrebbe ad una semiotica prestrutturale. Nella teoria semiotica di Eco, la ‘reversibilità totale’ tra i due piani costituisce allora un presupposto, programmaticamente esposto nel capitolo 0 del Trattato di semiotica generale, ed un concetto operativo, che è in grado di rendere conto della natura rizomatica dell’enciclopedia, operante ante litteram già nel Modello Q” (Ivi, p. 1). Di quest’ultimo parleremo tra un istante, ma in esso accade qualcosa di molto simile a ciò che interessa la reversibilità tra espressione e contenuto: ci riferiamo al rapporto type/token su cui si basa la competenza semantica minima di ogni parlante.

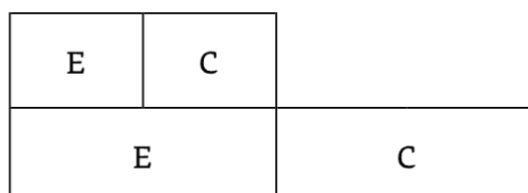


Figura 16: Espressione e contenuto di una funzione semiotica.¹⁴¹

Accade insomma qualcosa che cercavamo di mostrare nella comparazione tra l’approccio VADER e l’approccio RoBERTa nella *sentiment analysis*: da un lato, il primo, equivalenza, dall’altro, il secondo, inferenza. Da un lato s-codice, dall’altro Enciclopedia. Se applichiamo l’esempio di

grado immediatamente inferiore, e grazie alla somma delle dipendenze che queste parti di ordine immediatamente inferiore contraggono fra di loro”. Lo strutturalismo di Hjelmslev definisce la totalità non come composta da “cose”, ma da “rapporti” e le entità che costituiscono il rapporto non sono antecedenti al rapporto stesso. Espressione e contenuto sono i funtivi che contraggono una funzione – la funzione segnica - e che acquisiscono senso in ragione della loro dipendenza reciproca.

¹⁴⁰ Valle, A. (2003). Le due facce del senso. Note su espressione e contenuto. *Semiotiche*, 1, 2003, pp. 13–43.

¹⁴¹ In Paolucci (2010), p. 363.

semiotica “meta” che Paolucci e Valle riprendono da Hjelmslev ed Eco, possiamo forse dire che nel passaggio dal prompt al modello generativo siamo in presenza di un’occorrenza del tipo “semiotica meta”? Il prompt entra nel modello come funzione semiotica “standard” del linguaggio verbale, per trascrizione si tramuta in token, che a sua volta diviene espressione all’interno di una nuova funzione segnica, in cui sistema veicolante e sistema veicolato non preesistono l’atto stesso di instaurazione della funzione. E una volta all’interno del Transformer le operazioni tra matrici, vettori, tensori che si danno nel processo generativo e che rimodellano lo spazio latente in accordo a specifici attrattori di senso, si moltiplicano. Allo stesso modo, se assumiamo Hjelmslev, cosa che tenderemmo a fare, non possiamo parlare di un LLM come di un sistema monoplanare che opera attraverso la sola manipolazione del significante, cioè “Non ci può dunque essere, tranne che per un’artificiale separazione, un contenuto senza un’espressione, né un’espressione senza un contenuto” (Hjelmslev 1961, p. 53). In fondo, ci sembra così rispettato quel principio di *semiosi illimitata* che anima la vita dell’Enciclopedia echiana. E, localmente, la via interpretativa della *semiosi illimitata* è, nella nostra visione dell’ingegneria della macchina, imboccata a partire dal prompt umano, come in una passeggiata tra i paesaggi epigenetici di Waddington.

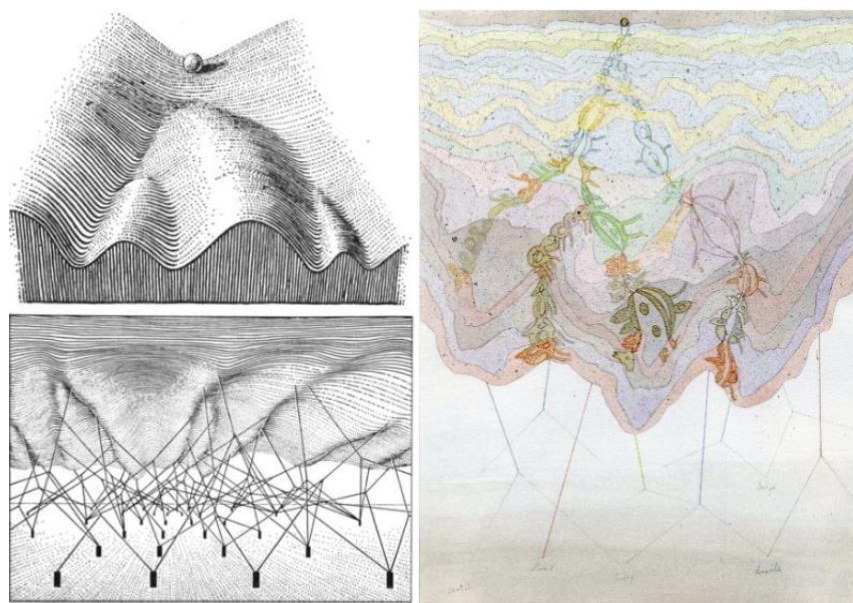


Figura 17: Paesaggi epigenetici e loro interpretazioni¹⁴²

¹⁴² In basso a sinistra, il sistema di interazioni alla base del paesaggio epigenetico in Waddington (1957), ripubblicato in Waddington, C. H. (2014). *The Strategy of the Genes*. Routledge. P. 36 <https://doi.org/10.4324/9781315765471>; in alto a sinistra, parte del paesaggio epigenetico. Leggiamo nella didascalia dell’immagine in basso a sinistra: “The pegs in the ground represent genes; the strings leading from them the chemical tendencies which the genes produce. The modelling of the epigenetic

Rispetto a quanto sostenuto sinora, questa metafora potrebbe apparire contraddittoria e lo è nella misura in cui il lavoro di Waddington si muove entro le strutture che definiscono il “campo morfogenetico degli embriologi”¹⁴³, come lo delimita René Thom che a questo si ispirava nell’elaborazione delle sue dinamiche qualitative. Per arrivare a sostenere che “il problema fondamentale della biologia è un problema di topologia, poiché la topologia è precisamente la struttura matematica che permette il passaggio dal locale al globale” (Ivi, p. 169)¹⁴⁴. Il metodo dei modelli locali di René Thom risuona fortemente con il modello di Enciclopedia teorizzato da Eco, poiché entrambi sanciscono il primato epistemologico del locale sul globale. Così lo spiega Thom: “Uno degli interessi essenziali del metodo dei modelli locali qui preconizzato è che esso non pregiudica in alcun modo la natura ultima della realtà; anche se quest’ultima dovesse rivelarsi all’analisi di una complessità da scoraggiare ogni descrizione, solo alcuni dei suoi aspetti entrerebbero alla fine nella descrizione macroscopica: quelli che determinano parametri <<osservabili>> del sistema. [...] Ad ogni sistema parziale di fenomeni relativamente indipendente dal mezzo che lo circonda, ci si sforzerà di associare un modello locale che renda conto qualitativamente e, nel migliore dei casi, quantitativamente, della fenomenologia del sistema. Ma non si può sperare a priori di integrare tutti questi modelli locali in una struttura globale; se fosse effettivamente possibile l’uomo potrebbe fondatamente dire di conoscere la *natura ultima della realtà*” (Ivi, p. 110). Nel 2015, Patrizia Violi dedica un articolo di *Semiotica* al rapporto tra globale e locale nella determinazione del significato delle parole, tracciando una traiettoria di fallimenti interni alle molteplici teorie che hanno, da Saussure in poi, cercato di definire quella che chiama “l’illusione del significante”: dai modelli universali chiusi (strutturalismo) alla possibilità di una descrizione delle regolarità di un’unità linguistica (modelli prototipici). Lo dicevamo già all’inizio, sembra non esserci spazio per una teoria

landscape, which slopes down from above one's head towards the distance, is controlled by the pull of these numerous guy-ropes which are ultimately anchored to the genes”. Leggiamo nella discalìa dell’immagine in alto a sinistra: “The path followed by the ball, as it rolls down towards the spectator, corresponds to the developmental history of a particular part of the egg. There is first an alternative, towards the right or the left. Along the former path, a second alternative is offered; along the path to the left, the main channel continues leftwards, but there is an alternative path which, however, can only be reached over a threshold” (p. 29). Siamo di fronte a un procedure per biforcazioni. A destra: Gemma Anderson, “Isomorphogenesis embedded in Waddington’s Epigenetic Landscape.” Watercolour on paper, 2016 in Anderson, G., Verd, B., & Jaeger, J. (2020). *Drawing to Extend Waddington’s Epigenetic Landscape*. Leonardo, 53(3), 256–262. https://doi.org/10.1162/leon_a_01738

¹⁴³ Thom, R. (1977), *Stabilité structurelle et morphogénèse. Essai d’une théorie générale des modèles*, InterEditions, Paris. Tr. It. *Stabilità strutturale e morfogenesi. Saggio di una teoria generale dei modelli*, Ghibli (2023).

¹⁴⁴ Menzioniamo per ragioni di continuità concettuale e per memoria storica, senza poterne dare descrizione, sempre sul tema Locale/globale la voce omonima dell’*Enciclopedia Einaudi*, VIII vol. Labirinto-Memoria, curata da Jean Peitot.

del significato, per un trattamento sistematico del lessico, se non nella forma di un *potenziale semantico* che si insinua a livello della svolta testuale, proprio nell'articolazione tra locale (il testo) e globale (l'enciclopedia).¹⁴⁵

Nel suo saggio *L'Enciclopedia: criticità e attualità*,¹⁴⁶ di due anni successivo, Violi ricostruisce con metodo archeologico l'evoluzione del concetto di Enciclopedia in Eco a partire dalle prime tracce antecedenti al Trattato di semiotica generale: "l'Enciclopedia sembra attraversare come una sorta di filo rosso tutta la produzione scientifica e letteraria di Eco. Elaborata in un primo tempo nel suo lavoro teorico, emigra poi nel mondo di finzione dei romanzi". (Ivi, p. 367). Questo filo rosso che tiene insieme produzione teorica e produzione narrativa, in Eco irreparabilmente connesse, accompagna una trasformazione dei concetti base della semiotica echiana, evidenzia Violi: alle rigide correlazioni della prima impostazione strutturalista si sostituisce una visione aperta, dinamica e potenzialmente illimitata dello *Spazio Semantico Globale*, potremmo aggiungere virtualmente e intensivamente illimitata, nella nostra interpretazione.

¹⁴⁵ Sintetizziamo di seguito i sette punti che Violi propone a partire dalla tensione tra locale e globale all'interno di una teoria del significato (Violi 2015, p. 100-102):

1. I termini non rimandano a significati fissi, ma a un potenziale semantico di natura enciclopedica (Hjelmslev: il purport del piano del contenuto).
2. Le parole sono attivatori di contesti e inferenze: sono dispositivi abducenti che creano le proprie condizioni di applicabilità, non sono solo entità passive che vengono determinate dal contesto. Se le parole fungono da istruzioni o da attrattori di contesto, il meccanismo alla base è il rapporto tra invariante (type) e variabilità (token), dove "we should think of actualized tokens through which the step-by-step configurations of semantic potential are specified".
3. Le descrizioni prototipiche, che catturano le regolarità ricorrenti, sono utili ma limitate: restano una "semiotica del senso comune", non una teoria del sistema semantico in senso forte.
4. Il dispositivo di conversione dal potenziale semantico al piano locale del sistema testuale è l'enunciazione.
5. L'enunciazione è atto di un soggetto incarnato che converte il potenziale globale nel senso locale. Qui fa emergere una duplice variante della prassi enunciativa: da un lato è dispositivo formale di soggettivazione (Benveniste), dall'altro riorganizza localmente il significato (per qualcuno, sotto qualche rispetto).
6. In riferimento alla teoria semiotica dei modi di esistenza, il potenziale semantico corrisponde al virtuale enciclopedico echiano. L'attualizzazione del potenziale è in carico all'enunciatore, mentre il modo realizzato si costituisce nel significato locale.
7. Il significato locale è una porzione riconfigurata e riproporzionata del potenziale semantico, avvicinandosi dice Violi alla sostanza in Hjelmslev, per il quale è connessa agli usi. Ciò che dà forma alla materia del contenuto non è uno schema formale, ma l'attività di enunciazione e del suo soggetto.

¹⁴⁶ Violi, P. (2017), *Encyclopedia: Criticality and actuality* in Beardsworth, S.G., Auxier, R.E. (2017), *The philosophy of Umberto Eco. The library of living philosophers. Volume XXXV*, Open Court, pp. 223-250. Tr. It. Violi, P. (2015), *L'Enciclopedia: criticità e attualità* in Lorusso A. M. (2021). a cura di, *La filosofia di Umberto Eco. ITA*. https://cris.unibo.it/handle/11585/802029_367-407, pp. 367-405.

Da una prospettiva geometrica l'enciclopedia ha la forma di un labirinto a rete "in cui ogni punto può essere connesso con qualsiasi altro punto", con dimensionalità infinita e priva di tratti topologici che identifichino un interno e un esterno.

"Poiché ogni suo punto può essere connesso con qualsiasi altro punto, e il processo di connessione è anche un processo continuo di correzione delle connessioni, la sua struttura sarebbe sempre diversa da quella che era un istante prima, e ogni volta si potrebbe percorrerlo secondo linee diverse. Quindi chi vi viaggia deve anche imparare a correggere di continuo l'immagine che si fa di esso, sia questa una concreta immagine di una sua sezione (locale), sia essa l'immagine regolatrice e ipotetica che concerne la sua struttura globale (inconoscibile, sia per ragioni sincroniche che per ragioni diacroniche)" (ivi, p.74).

La metafora scelta da Eco (1984, 2007) per visualizzare il modello-rete dell'enciclopedia, è il rizoma per la sua natura non gerarchizzata, "smontabile e reversibile, suscettibile di modificazioni", sempre aperta e contraddittoria¹⁴⁷, dove possono verificarsi processi di loop e dove l'algoritmo di percorrenza è miope, cioè costruisce una visione, una descrizione valida localmente e congetturalmente, che è solo "un'ipotesi circa la globalità della rete".

Le semantiche enciclopediche sono state oggetto di diversi tentativi di rappresentazione, ribadiamo sempre di natura locale, ma suscettibili di "arricchimento progressivo e potenzialmente senza limiti". Eco stesso ha proposto un modello a selezione contestuale e circostanziale, nel quale "si specificano i sensi che un dato lessema assume in diversi contesti" (Eco 2007, p. 76). Se un tale modello consente di far funzionare il principio peirciano di interpretazione e quindi garantire un processo semiotico che si autosostiene, esso manca ancora del ricorso a struttura a rete.

Nota infatti Eco che "È nel campo dell'Intelligenza Artificiale che appaiono rappresentazioni a *frames e scripts o sceneggiature*, che registrano ogni stadio di una sequenza di eventi tipici – tutti modelli che hanno avuto fortuna nell'ambito dell'Intelligenza Artificiale dove, per permettere a un computer di comprendere un testo e trarne inferenze occorre fornirgli tutte le competenze di cui (anche senza rendersene conto) è fornito un essere umano medio" (Ivi, p. 77).

¹⁴⁷ Per la compatibilità della presenza di occorrenze contraddittorie nella forma di determinazioni negative, rispetto al concetto di *virtuale* nell'accezione deleuziana si veda il paragrafo successivo e si veda Paolucci, C. (2010), in particolare il capitolo 3 "Una nuova teoria della differenza".

Eco attribuisce a Ross Quillian¹⁴⁸ la formulazione di un modello di natura semantica in forma di rete, strutturato come “un labirinto di nodi interconnessi”. La particolarità di questo modello, che Eco nel *Trattato di semiotica generale* chiama Modello Q è che si fonda sulla relazione type/token in base alla quale ciascun nodo costituisce il “capostipite” (type) di una serie di altri nodi (tokens) rappresentanti i termini che lo definiscono e che, a loro volta, possono divenire type di un’altra serie di tokens. Questi termini “definienti” vanno a comporre l’universo degli interpretanti di quel determinato concetto type, attraverso una rete interconnessa che non ha un centro e una periferia assoluti, ma solo relativi.

A fronte dell’irrappresentabilità grafica bidimensionale di un tale modello nella sua complessità, Eco ci dice che “Esso dovrebbe apparire come una sorta di rete polidimensionale, dotata di proprietà topologiche, dove i percorsi si accorciano e si allungano e ogni termine acquista vicinanze con altri, attraverso scorciatoie e contatti immediati, rimanendo nel contempo legato a tutti gli altri secondo relazioni storicamente mutevoli. È stato detto che, se si assume una nozione massimale di competenza intorno al mondo, allora il significato di un termine sarebbe costituito da tutti gli enunciati veri in cui esso è apparso o potrebbe apparire” (Ivi, p. 78).

3.4 Passeggiate inferenziali e ottimizzazioni

¹⁴⁸ Riprendiamo da Quillian, R. (1961), A revised design for an understanding machine, in *Mechanical Translation*, Vol.7, no.1, July 1962, p. 19: “Let us approach the problem by considering first the totality of information on the basis of which a person acts at any particular moment, including both the information which he is consciously aware of having, and that which he has but is in greater or lesser degree not conscious of having. We shall think of this information as flowing into whatever center or centers there may be in the person which direct his action. It flows in from exteroceptors connected to the outside world, from interoceptors and proprioceptors describing conditions within his body, and also from his “memory”. The information from “memory” provides him with such notions as that of a constant, expanded space, in which objects are located. It continuously enlarges his perceptual world to include some “knowledge” of things which he is not actually sensing at the moment. At any one instant these several flows of information combine to produce a broad, rushing stream of input to what for convenience we will simply call the person’s “action direction center”. Now some of this information input – if not all of it – becomes transformed into “meaningful” information before or as it reaches the person’s action direction center. We may ask: What is the nature of the transformation it undergoes in so changing from raw sensory input into meaningful information? It has already been realized by at least some writers that the operation which is performed on a bit of sensory input as it becomes meaningful perception is one of its being related to other information. This process of “becoming related” to other information seems to me to be usefully viewed as two simultaneously occurring processes. First, the bit of information may be said to be combined with other information which is flowing in at approximately the same time, thus creating the celebrated “gestalt” of perception. Secondly, the information formed into such gestalts can be considered to be compared to yet other information which in general is not part of that flowing into the action direction center at that moment.

Questo tipo di percorrenze interne ai paesaggi semantici enciclopedici di un lessema, che in parte abbiamo cercato di visualizzare attraverso i creodi di Waddington, richiama inoltre le metaeuristiche utilizzate per addestrare le reti neurali artificiali, dove l'individuazione dei pesi migliori della rete (ossia i luoghi di concentrazione della conoscenza appresa dal modello) si presenta nella forma di un problema di ottimizzazione. Abbiamo in precedenza visto che i) obiettivo del training è ottenere l'output desiderato nell'ultimo layer della rete e ii) questo obiettivo è performato ricalibrando i pesi che connettono i neuroni di un layer con quelli del layer precedente. In compiti di classificazione, tendenzialmente ciò avviene in ambiente di training supervisionato e l'algoritmo maggiormente utilizzato in fase di addestramento è l'algoritmo di back-propagation che “adapts weights according to the mean squared error at the end nodes and which is then back-propagated to the previous layers. BP is essentially a gradient-based algorithm” (Pedram et al. 2022)¹⁴⁹. Secondo gli/le autori/autrici del paper appena citato, la retro-propagazione ha due inconvenienti: da un lato quello di restare bloccata nei punti di ottimo locale¹⁵⁰ e dall'altro quello della sensibilità ai pesi inizialmente appresi.

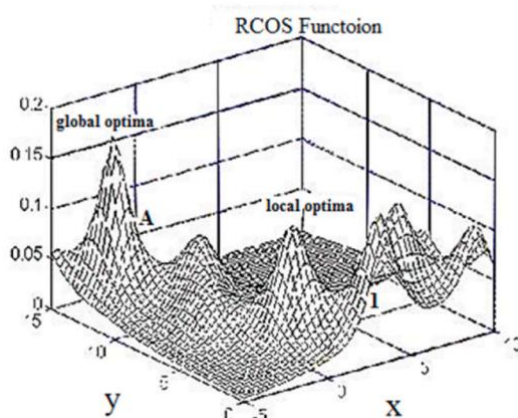


Figura 18: Punti di ottimo locale e globale¹⁵¹

¹⁴⁹ Pedram, M., Mousavirad, S.J., Schaefer, G. (2022). Training Neural Networks with Lévy Flight Distribution Algorithm. In: Kim, J.H., Deep, K., Geem, Z.W., Sadollah, A., Yadav, A. (2022) *Proceedings of 7th International Conference on Harmony Search, Soft Computing and Applications. Lecture Notes on Data Engineering and Communications Technologies*, vol 140. Springer, Singapore. https://doi-org.ezproxy.unibo.it/10.1007/978-981-19-2948-9_10

¹⁵⁰ Un punto di ottimo locale (massimo/minimo) di una funzione può essere definito come un elemento in cui, entro una certa area vicina, tutti i valori di una funzione sono rispettivamente più grandi/più piccoli o uguali rispetto al valore rappresentato dall'elemento stesso. Il concetto di ottimo locale o globale si inserisce nell'ambito della matematica che studia i problemi di ottimizzazione, ossia quei problemi che mirano a individuare la minimizzazione o massimizzazione di una funzione a valori reali su un insieme specificato. Per una rassegna degli algoritmi più aggiornati relativi ai problemi di ottimizzazione si veda: Mohammadi, A., & Sheikholeslam, F. (2023). Intelligent optimization: Literature review and state-of-the-art algorithms (1965–2022). *Engineering Applications of Artificial Intelligence*, 126, 106959. <https://doi.org/10.1016/j.engappai.2023.106959>

¹⁵¹ Mohammadi, A., & Sheikholeslam, F. (2023).

Le metaeuristiche, derivate spesso matematicamente dai comportamenti motori di popolazioni – animali ad esempio: formiche, balene, predatori marini e così via – sono considerate più utili in compiti di esplorazione di spazi di ricerca (*search space*) ampi – ossia l’insieme di tutte le potenziali soluzioni tra le quali la soluzione ottimale ad un problema possa essere identificata – non limitandosi all’ottimo locale.¹⁵² Pedram et al., usano la distribuzione del volo di Lévy, introdotta come algoritmo nel 2020 da Houssein et al.¹⁵³, per addestrare una rete neurale artificiale basata sul MultiLayer Perceptrons (MLP) e confrontano i risultati con altri metodi metaeuristici. In particolare, i voli, chiamati anche in taluni casi i cammini, di Lévy sono processi stocastici appartenenti alla classe dei processi di diffusione descritti dai cammini aleatori (random walks), di cui un esempio è il moto Browniano.

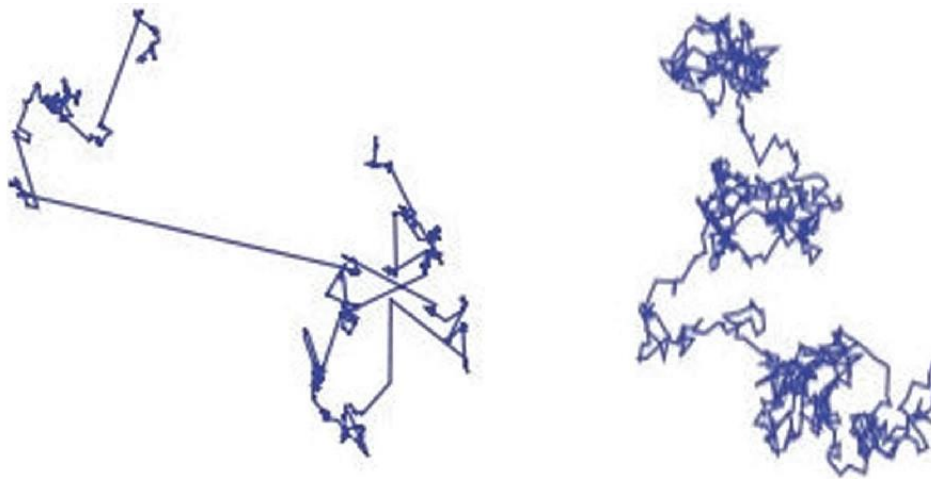


Figura 19: Cammino di Lévy a sinistra, moto Browniano a destra.¹⁵⁴

Senza entrare nei dettagli tecnici dei due processi diffusivi illustrati in Figura 19, è immediatamente evidente il diverso pattern di movimento all’interno dello spazio di ricerca: il cammino di Lévy, infatti, combina passi molto corti nella maggior parte casi con salti lunghi più sporadici; invece, nel moto Browniano i passi sono continui, corti e locali. Dunque, il primo tende a

¹⁵² Per una review comprensiva dei metodi di ottimizzazione attraverso meta-euristiche applicate all’addestramento di reti neurali artificiali si veda: Kaveh, M., Mesgari, M.S. (2023), Application of Meta-Heuristic Algorithms for Training Neural Networks and Deep Learning Architectures: A Comprehensive Review. *Neural Process Lett* 55, 4519–4622 (2023). <https://doi.org.ezproxy.unibo.it/10.1007/s11063-022-11055-6>

¹⁵³ Houssein, E. H., Saad, M. R., Hashim, F. A., Shaban, H., & Hassaballah, M. (2020). Lévy flight distribution: A new metaheuristic algorithm for solving engineering optimization problems. *Engineering Applications of Artificial Intelligence*, 94, 103731. <https://doi.org/10.1016/j.engappai.2020.103731>

¹⁵⁴ In Pedram et al. 2022.

evitare fenomeni di stagnazione a discapito della stabilità della soluzione rendendosi una metaeuristica adatta a un'esplorazione globale del *search space*, invece il secondo garantisce una convergenza stabile con potenziale rischio di intrappolamento nei minimi locali, prestandosi maggiormente a un affinamento fine e a una ricerca locale. La pertinenza di questo affondo tecnico rispetto al nostro discorso sull'irrappresentabilità globale dell'Enciclopedia rimanda alla natura dello spazio latente che abbiamo detto rappresentare il virtuale intensivo dalla cui attivazione è possibile generare nuove forme, ossia nuovi enunciati. Il fatto che le metaeuristiche servano in fase di training per addestrare il modello a leggere le relazioni tra i dati implica poter prendere in considerazione il modo operativo in cui il modello costruisce la rappresentazione dei pattern e delle dipendenze tra le unità culturali che costituiscono gli archivi dei modelli generativi.

In questo senso, il nostro procedere per analogia ci porta alla conclusione di questa panoramica sull'Enciclopedia echiana, mettendoci nella condizione di fare due operazioni: i) la prima è quella di promuovere un passaggio dello status dell'idea di Enciclopedia da ipotesi regolativa a ipotesi operativa; la logica di funzionamento dell'Enciclopedia come descritta nelle pagine precedenti è perfettamente funzionante su scala (ancora) ridotta, riteniamo, all'interno dei grandi modelli di linguaggio; ii) la seconda è di riaffermare quanto sostenuto nell'incipit introduttivo della tesi: il modello di IA generativa, in fase di training, apprende un modello del mondo mediato dai testi di addestramento che funge da virtuale intensivo per la generazione di nuove forme, prodotte dall'assemblaggio di operatori radicalmente eterogenei: ad un primo livello, umano e macchinico, a loro volta inglobanti istanze enuncianti eterogenee.

Prima di passare a quest'ultimo aspetto che abbiamo definito quello di un'IA per sé, nel capitolo 6, vorremmo soffermarci sul carattere intensivo del virtuale enciclopedico, anche per fornire maggiore consistenza teorica alla nostra tesi.

4.0 L'Enciclopedia echiana come virtuale intensivo

In apertura dicevamo che dal nostro punto di vista un modello artificiale generativo ha bisogno di qualcosa come “un’Enciclopedia” per poter generare nuovi enunciati; abbiamo inoltre sottolineato che l’Enciclopedia non è identificabile in via esclusiva con i dataset/archivi con cui i modelli vengono addestrati; i dati utilizzati nel training, presi in se stessi, non darebbero origine a nuove occorrenze, a nuovi enunciati, poiché gli enunciati che popolano dell’Enciclopedia derivano sì da occorrenze estese, ma ciò che l’Enciclopedia come ipotesi operativa deve conservare sono le relazioni differenziali tra gli enunciati, la cui esplorazione deve essere retta dal principio di interpretazione peirciano.

Eco usa l’immagine del rizoma, proposta da Deleuze e Guattari in *Mille Piani*, per rendere figura la geometria dell’Enciclopedia, come abbiamo visto. Ma è nella teoria dell’enunciazione formulata da Paolucci che troviamo un’idea “chiara e distinta” del modo in cui le grandezze semiotiche che compongono la virtualità enciclopedica vengono convocate all’interno degli enunciati. L’enunciazione impersonale ed evenemenziale, una teoria che per dichiarazione del suo stesso autore risente fortemente del pensiero di Deleuze, vuole essere una “topologia della soggettività nel linguaggio”, allontanandosi da un’ontologia del soggetto come condizione dell’enunciazione. Uscire dall’ontologia significa per Paolucci operare in diverse direzioni: “uscire dalla *soggettività* del locutore, uscire dall’*intersoggettività* del suo interlocutore, uscire infine dall’*oggettività* dell’enunciato o dell’oggetto del discorso, “ciò di cui si parla” con il suo statuto delocutivo” (Paolucci 2020, p. 20). Ciò che caratterizza la vita di questo “piccolo dramma” – come Paolucci definisce l’enunciazione, riprendendo Tesnière – sono dei movimenti di passaggio tra modi di esistenza, tra i quali l’enunciazione consente di mediare. I modi di esistenza semiotici (Fontanille & Zilberberg 1998; Fontanille 2003) vengono da Paolucci coniugati con l’idea di Enciclopedia per rendere conto di come nell’atto di enunciazione si diano contemporaneamente le istanze che enunciano e l’enunciato stesso, aprendo così spazio “a forme di vita non necessariamente linguistiche”.

L’abbiamo in parte anticipato nel paragrafo precedente, ma riprendiamo un passaggio di *Persona* per renderne maggiormente ragione: “Per essere attivata concretamente, l’enciclopedia deve forzatamente essere ritagliata, ma ogni ritaglio locale serve a stabilizzare e regolarizzare il significato attraverso un modello esplicativo che a livello globale è invece costitutivamente instabile, dal momento che, come notava lo stesso Eco, esso è divenuto diverso da sé già nel tempo che serve per descriverlo. L’essenza dell’enciclopedia è questa: la doppia natura di modello globale instabile costituito da domini locali stabili che definiscono regolarità. È una specie di patchwork, formato da brandelli semiotici essenzialmente eterogenei che ogni atto di enunciazione prova a ricucire a suo

modo. Data la corrispondenza tra enciclopedia e rizoma stabilita da Eco (1983, 1984), una porzione locale di enciclopedia, un suo “ritaglio”, ne sarà un “piano” (Paolucci 2020, p. 108).¹⁵⁵

Poniamo l’accento sui “brandelli semiotici eterogenei” che ogni atto di enunciazione prova ad assemblare: da un lato abbiamo “una congerie di enunciati già enunciati (realizzati)” dall’altro “lo sfondo di nuove enunciazioni (virtuale)”; il condizionamento è reciproco e la codeterminazione di modi di esistenza eterogenei è imprescindibile. Così, nel nostro modello artificiale generativo abbiamo: i dataset/archivi che costituiscono “la congerie di enunciati già enunciati” e il virtuale (lo sfondo) è indefinibile in termini di una geometria estesa, come lo era il rapporto tra globale e locale o tra geometrie riemanniane e geometrie euclidee descritte da Petitot nella voce dell’Enciclopedia Einaudi citata poco sopra.¹⁵⁶ Tant’è che le metafore utilizzate per esprimere ciò che accade all’interno di una deep neural network (rete neurale profonda) variano dalla black box (scatola nera) alla mera probabilità statistica. La prima metafora apre l’abisso dell’incontrollabile, la seconda apre alla velleità dell’irriproducibile a livello di formalizzazione. Entrambe le metafore fanno appello alla controllabilità dei sistemi generativi.

4.1 Dei virtuali molteplici o delle molteplicità del virtuale.

Sempre da una prospettiva semiotica, Fontanille et Zilberberg riconoscono un fatto critico generale di cui bisogna tenere conto: “à tout moment de l’évolution d’une culture et des discours qui la constituent, en tout point de sa diffusion, cohabitent au moins deux types de grandeurs: celles engendrées à partir du système, et celles fixées par l’usage” (Fontanille et Zilberberg 1998, p. 129). Alle prime (in potenza) appartengono i modi di esistenza “virtualizzato” e “potenzializzato”, al secondo (in atto) i modi “attualizzato” e “realizzato”. Quando si arriva alla caratterizzazione dei singoli modi, rispetto al modo virtualizzato viene riferito quanto segue: “les formes sémio-narratives

¹⁵⁵ Su questo si veda anche Paolucci 2010, p. 57: “Deleuze e Guattari (1980) chiamavano rizoma l’unione di uno spazio liscio e di uno spazio striato. Non è dunque un caso se, seguendo Petitot, si era potuta riconoscere una struttura partecipativa come costitutiva dell’enciclopedia. Se infatti l’enciclopedia è un rizoma, e se un rizoma è l’unione di uno spazio liscio e di uno spazio striato, l’enciclopedia ha per essenza una natura partecipativa, dal momento che uno spazio striato definisce una zona precisa (intensiva) ottenuta per divisione di uno spazio continuo in domini attraverso un sistema di frontiere, là dove uno spazio liscio definisce l’annullamento stesso di questa divisione, che dà vita a fenomeni di oltrepassamento delle frontiere (estensivi). C’è qui la possibilità di distinguere molto chiaramente il rizoma dalla struttura. La struttura è uno spazio striato o l’unione di più spazi striati [...]. Al contrario, il rizoma è l’unione di almeno uno spazio liscio e di almeno uno spazio striato, e cioè un “incollamento” di strutture e di zone vaghe, un “incollamento” di sistemi descrivibili attraverso un sistema di opposizioni locali precise e determinate (intensive) e di zone in cui i sistemi invece si sovrappongono, si compenetrano e vedono definita la loro stessa identità dalla possibilità di un annullamento della loro stessa struttura su un altro livello (estensive)”.

¹⁵⁶ Cfr. ancora Paolucci 2010, pp. 53-54.

(le système) constituent la compétence énonciative virtuelle” e, poco più avanti, “le mode virtualisé et le mode potentialisé correspondent tout deux à l’état latent des formes disponibles, au langage ‘en puissance’ selon Guillaume, au ‘système’ selon Hjelmslev” (Ivi, pp. 129-130).

A sua volta Hjelmslev nei *Fondamenti* così definisce la virtualità: “si dice che una classe è virtuale se di essa non si può fare l’oggetto di un’analisi particolare” in contrapposizione alla realizzazione: “si dice che una classe è realizzata se di essa si può fare l’oggetto di un’analisi particolare” (Hjelmslev 1961, p. 147). Senza entrare nel merito della definizione di analisi particolare, che ha una serie di implicazioni teoriche, ciò che ci sembra rilevante al fine di determinare cosa sia la virtualità è la sua definizione per via negativa: non ci viene detto cosa sia il virtuale, viene però detto che è effetto di ciò che non può essere fatto oggetto di analisi particolare, ossia di una o più partizioni. Al di fuori del suo sistema di definizioni e a proposito del rapporto tra lingua e testo, cioè tra sistema e processo, Hjelmslev ci indica la relazione reciproca tra virtuale e realizzato: “È dunque impossibile avere un testo senza una lingua ad esso soggiacente. D’altra parte, si può avere una lingua senza un testo costruito in tale lingua. Questo significa che la lingua in questione è prevista dalla teoria linguistica come un sistema possibile, ma che nessun processo appartenente a tale sistema è presente in maniera *realizzata*. Il processo testuale è allora *virtuale*” (Ivi, p. 44). La lingua, intesa come sistema e come teorizzata dal linguista, non può essere fatta oggetto di analisi se non a partire da testi “che esistano naturalmente o che siano costruiti dal teorico sulla base del sistema”; in questo caso il testo è virtuale come possibilità della lingua. A sua volta però il sistema della lingua deve essere presupposto come realizzato, anche nel caso di testo virtuale, anche se non può essere fatto oggetto di analisi particolare. Vedremo che questo intreccio tra virtuale, reale e possibile assume tutt’altra configurazione in Deleuze, il quale, nella sua collaborazione con Guattari, era solito riferirsi a Hjelmslev, come “il geologo danese spinoziano... il tenebroso discendente di Amleto” (Deleuze & Guattari 1980, p. 68). Ma andiamo per gradi e torniamo a Paolucci (2020), dove troviamo già mobilitate le categorie teoriche a noi necessarie per parlare dell’Enciclopedia come un virtuale intensivo.

“Al fine di farla funzionare, la teoria dell’enunciazione dovrà allora determinare qual è la sua “scena predicativa”, qual è il suo “piccolo dramma”, qual è cioè il posto del soggetto, qual è il posto della società, qual è il posto dell’enunciato, qual è il posto dello schema, delle norme e degli usi, quale posto hanno insomma le varie istanze enuncianti all’interno della “scena” dell’atto di enunciazione” (Ivi, p. 101). E di seguito, identificando nell’immagine e semiotica e matematica dell’eterogenesi differenziale che Sarti, Citti e Piotrowski (2019, 2022) propongono ispirandosi a Deleuze, descrive così il virtuale enciclopedico:

“data la coesistenza virtuale di differenti istanze eterogenee quali schema, norme, usi, apprezzamenti collettivi ed enunciati (“assemblaggio di operatori eterogenei”), l’enunciazione è una forma di “eterogenesi differenziale”, in cui si passa dal virtuale all’attuale. Il piano virtuale (norme, abiti, schemi ecc.) è intensivo e “non può essere percepito, dal momento che non appartiene al piano fenomenico” (Sarti, Citti e Piotrowski, 2019, p. 3), ma la sua integrazione e attualizzazione attraverso l’atto dà vita a “forme, percezioni e morfologie estese”, che sono invece attuali e sono percepibili all’interno dell’enunciato”. (Paolucci 2020, p. 101).

Il piano intensivo di cui parla Paolucci, a partire dal lavoro di Sarti, Citti e Piotrowski è il nome con cui Deleuze ribattezzò il pre-individuale simondoniano, quando nel 1966 scrisse una – breve, ma *intensa* – review della tesi di dottorato di Simondon dal titolo “L’individu et sa genèse physico-biologique”.¹⁵⁷ Sintetizzando molto per ragione di spazi, Simondon contesta la traiettoria di analisi che dall’individuo – variamente considerato in accordo ad un modello sostanzialista/monista o in accordo a un modello ilomorfico/dualista – ricerca la sua caratterizzazione nel principio di individuazione, oscurando l’operazione di individuazione, ossia evitando di ricercarne la sua spiegazione. Dice Simondon: “Vorremmo dimostrare che occorre operare una svolta nella ricerca del principio d’individuazione, considerando come primaria l’operazione d’individuazione a partire dalla quale l’individuo acquisisce la sua esistenza e della quale riflette lo sviluppo, il regime e, infine, le modalità nei suoi caratteri propri. L’individuo sarebbe così concepito come realtà relativa, come una certa fase dell’essere che presuppone una realtà preindividuale e che, anche dopo l’individuazione, non esiste di per sé, poiché, a sua volta, l’individuazione non esaurisce di colpo i potenziali della realtà preindividuale. D’altra parte, ciò che l’individuazione fa apparire non è solo l’individuo bensì la coppia individuo-ambiente” (Simondon 1964, p. 33)¹⁵⁸. Ciò che viene dunque teorizzato, rispetto alle analisi classiche sul rapporto tra essere e divenire nella genesi dell’individuo sul piano ontogenetico è l’esistenza di una fase dell’essere definita pre-individuale, cioè una fase dell’essere in cui non sussistono fasi, per cui l’individuazione corrisponderebbe invece alla comparsa delle fasi dell’essere. Per poter comprendere questo piano pre-individuale occorre ripensare all’individuazione dalla

¹⁵⁷ Deleuze, G. (1966). *Review of Gilbert Simondon's L'individu et sa genèse physico-biologique*. Pli, 12(2001), 43-49. Come leggiamo nell’avvertenza dell’editore Millon: “La thèse de Doctorat de Gilbert Simondon intitulée: *L’Individuation à la Lumière des Notions de Forme et d’Information*, a paru partiellement en deux ouvrages séparés aux Presses Universitaires de France, (Collection «Epiméthée», 1964) sous le titre: *L’Individu et sa Genèse Physico-Biologique*. Et aux Editions Aubier: (Res L’Invention Philosophique, 1989) sous le titre: *L’Individuation Psychique et Collective*” in Simondon, G. (1995). *L’individu et sa genèse physico-biologique*. J. Millon, p. 7.

¹⁵⁸ Simondon, G. (1964), *L’Individuation à la Lumière des Notions de Forme et d’Information*, Presses Universitaires de France; tr. it. *L’individuazione alla luce delle nozioni di forma e d’informazione*, Mimesis (2011).

prospettiva di un equilibrio metastabile, un sistema di potenziali non attualizzati, cioè uno stato di “sovrasaturazione iniziale dell’essere senza divenire e omogeneo, che solo successivamente si struttura e diviene, consentendo la comparsa dell’individuo e dell’ambiente secondo il divenire che, a sua volta, consiste in una risoluzione delle tensioni primigenie ed in una conservazione di queste tensioni sotto forma di struttura” (Ivi, p. 34). Il pre-individuale, il sistema metastabile è caratterizzato dalla presenza di due ordini di grandezza incapaci di comunicare tra loro, dunque ordini di grandezza privi di mediazioni, immersi in una soluzione sovrasatura, che “è più di unità e più di identità”. Il preindividuale è lo stato di apparizione della complementarità tra concetti che nella microfisica appaiono a coppie: onda-corpuscolo, materia ed energia. Nel dominio del vivente, l’individuazione costituisce un’origine assoluta e uno sdoppiamento: da un lato la dualità individuo-ambiente, dall’altro “il vivente conserva in se stesso un’attività di individuazione permanente” (Ivi, p. 37).

Deleuze rilegge questo quadro nel suo commento alla tesi di Simondon, dichiarando che la differenza di potenziale che abita il sistema metastabile, il potenziale di energia distribuito entro i limiti del sistema, possa essere “reconciled with a theory of intensive quantities, since each intensive quantity is a difference in itself. An intensive quantity comprises a difference in itself, contains factors of the type E-E', to infinity, and is first established between disparate levels, heterogeneous orders that will only enter into communication later, in extension. Like any metastable system, it is a structure (not yet a synthesis) of the heterogeneous” (Deleuze 1966, p. 44). La teoria delle quantità intensive rimanda in Deleuze alla critica della riduzione kantiana dello spazio e del tempo a delle estensioni geometriche e all’attribuzione della quantità intensiva alla materia che occupa uno spazio, un’estensione secondo gradi differenti. È lo spazio in quanto intuizione pura ad essere quantità intensiva, non in relazione ai concetti dell’intelletto, ma alla capacità dell’Idea di determinare nell’estensione l’attualizzazione dei legami ideali, cioè i rapporti differenziali contenuti nell’Idea. La quantità intensiva è la differenza in sé, il diseguale in sé; afferma la differenza, cioè fa della differenza oggetto di affermazione; infine l’intensità, la quantità intensiva è implicata in sé e involupata: non è divisibile come la quantità estensiva (temperatura e velocità non sono composte da differenze dello stesso tipo – unità – ma implicano serie di termini eterogenei e cambi di natura) e non è indivisibile come la qualità poiché la differenza di qualità non è solo una differenza di natura, ma come in un gioco di prestigio qualcosa recupera quello che è perso da qualcos’altro (se qualcosa è più bello e qualcos’altro è più brutto). In sintesi, come avrebbe detto ChatGPT 3.5, “l’energia in generale o la quantità intensiva è lo *spatium*, teatro di ogni metamorfosi, differenza in sé comprendente tutti i propri

gradi nella produzione di ciascuno di loro. In questo senso l'energia, la quantità intensiva, è un principio trascendentale e non un concetto scientifico” (Deleuze 1968, p. 387).¹⁵⁹

Continuiamo a seguire Deleuze, per poi tirare alcune conclusioni sul rapporto tra Enciclopedia, virtuale, intensivo e modelli generativi. In *Differenza e ripetizione* Deleuze propone, inoltre, una prima trattazione più sistematica del virtuale nella sezione “La ripetizione per sé”, dove i rapporti di temporalità tra presente, passato e futuro sono letti alla luce delle tre sintesi, rispettivamente come *abitudine* (presente vivente), *memoria* (passato puro) e “*forma vuota del tempo*” (eterno ritorno, il ripetuto). Perché proprio qui? Perché a livello di instaurazione di molteplici Io locali nelle sintesi passive – di contemplazione o contrazione, le chiama Deleuze – si ha uno sdoppiamento di oggetti: “Da una parte, si stabilisce sulla fondazione delle sintesi passive una sintesi attiva, che consiste nel riferire l'eccitazione legata a un *oggetto posto come reale* e come termine delle nostre azioni (sintesi di riconoscimento fondata sulla sintesi passiva di riproduzione), ove la prova di realtà in una relazione detta «oggettuale» definisce la sintesi attiva, e per l'appunto, secondo il principio di realtà l'Io tende ad «attivarsi», a unificarsi attivamente, a raccogliere tutti i suoi piccoli Io passivi componenti e contemplanti, e a distinguersi topicamente dall'Es. Gli Io passivi erano già integrazioni, ma, come dicono i matematici, integrazioni soltanto locali, mentre l'Io attivo è tentata integrazione globale” (Ivi, p. 162). Dall'altra – e Deleuze porta l'esempio del bambino che impara a camminare – si costituisce un altro “oggetto o fuoco virtuale che regola e compensa gli insuccessi della sua attività reale”. A questa dualità oggettuale del costituirsi dell'Io, tra sintesi passive e sintesi attiva, corrispondo due serie correlative. Qui il discorso si sposta sulla differenziazione delle pulsioni di conservazione (inseparabili dalla fondazione della sintesi attiva e dai rapporti con l'oggetto reale) e delle pulsioni sessuali (inseparabili dall'approfondimento della sintesi passiva e dalla costituzione dei fuochi/oggetti virtuali). Dice Deleuze:

“Che le due serie non esistano l'una senza l'altra, sta a significare che non sono soltanto complementari, ma si scambiano e si alimentano reciprocamente in virtù della loro dissomiglianza o differenza di natura. Si constata così che *i virtuali sono prelevati sulla serie dei reali, e che sono incorporati nella serie dei reali*. Questo prelievo implica prima di tutto un isolamento o una sospensione, che paralizza il reale al fine di estrarne un atteggiamento, un aspetto o una parte. Ma tale isolamento è qualitativo, non consiste semplicemente nel sottrarre una parte dell'oggetto reale, in quanto la parte sottratta acquista una nuova natura funzionante come oggetto virtuale. Questo oggetto virtuale è un oggetto parziale, non solo perché manca di una parte rimasta nel reale, ma in sé e per sé,

¹⁵⁹ Deleuze, G. (1968). *Différence et répétition*, Paris, Presses Universitaires de France; tr. it. *Differenza e ripetizione*, Il Mulino (1971).

in quanto si sfalda, si sdoppia in due parti virtuali di cui l'una manca sempre all'altra. In altre parole, il virtuale non è subordinato al carattere globale che involge gli oggetti reali, dato che non solo per la sua origine, ma nella sua propria natura, il virtuale è brandello, frammento, spoglia, e non rispetta la propria identità” (Ivi, p. 165). Quasi per paradosso, Deleuze sembra rovesciare il rapporto tra globale/locale, ascrivendo la parzialità all’oggetto virtuale in rapporto all’oggetto reale globale. Tuttavia, come vedremo tra un’istante, virtuale e reale non sono i poli oppositivi in cui si articola la differenza e che, nella dinamica morfogenetica, determinano il divenire individuale. Deleuze sa di dover chiarire questo apparente paradosso.

Nel passaggio successivo, l’oggetto virtuale diviene passato, sul modello bergsoniano di un mondo a due fuochi, uno reale e l’altro virtuale, da cui derivano rispettivamente “la serie della immagini-percezione” e “la serie delle immagini-ricordi”. In questo passaggio, l’oggetto virtuale è “un lembo di passato puro”, un brandello, un frammento di passato puro. L’oggetto virtuale è una mancanza che non è negazione, ma è presenza di un’assenza.

In *Sintesi «ideale» della differenza*, il virtuale ricompare nel rapporto tra idea e concetto, o più propriamente: “L’Idea si presenta dunque sotto tre momenti: indeterminata nel suo oggetto, determinabile in rapporto agli oggetti dell’esperienza, e infine portatrice dell’ideale di una determinazione infinita in rapporto ai concetti dell’intelletto” (Ivi, p. 275-276). Nello stesso capitolo Deleuze pone il problema del differenziale, dell’infinitamente piccolo, poiché “l’Idea ha per oggetto il rapporto differenziale: essa integra allora la variazione, non più certamente come determinazione variabile di un rapporto che si suppone costante («variabilità»), ma viceversa come grado di variazione del rapporto stesso («varietà»), a cui corrisponde ad esempio la serie qualificata delle curve. Se l’Idea elimina la variabilità, questo avviene a vantaggio di ciò che deve denominarsi varietà o molteplicità” (Ivi, p.281). Qui il virtuale, la virtualità ricompare come una del tre condizioni che consentono l’emersione dell’Idea, come indeterminato:

- i) gli elementi della molteplicità non posseggono forma sensibile, significato concettuale, funzione specifica ed esistenza attuale. Non hanno dunque alcuna identità o posizione, sono differenziali. Essi sono inseparabili da un potenziale o da una virtualità. In virtù della loro indeterminazione rendono possibile la manifestazione della differenza;
- ii) tali elementi si determinano reciprocamente sulla base di rapporti di dipendenza, ossia connessioni ideali, non localizzabili all’interno della molteplicità che si definisce in maniera intrinseca, senza uscire da sé e senza ricercare l’uniformità. Proiettate nello spazio-tempo, le relazioni conservano la molteplicità, ma perdono l’interiorità; i concetti dell’intelletto viceversa, conservano l’interiorità, ma perdono la molteplicità;

- iii) le connessioni (i rapporti differenziali) si attualizzano in relazioni spazio-temporali, gli elementi si incarnano in atto in “termini e forme variate”.

L’Idea diviene così struttura, “una molteplicità interna, cioè un sistema di connessioni molteplici non localizzabili tra elementi differenziali, che s’incarna in relazioni reali e in termini attuali” (Ivi, p. 297). Il movimento di genesi dell’Idea non procede di attuale in attuale, ma dal virtuale alla sua attualizzazione, ossia “dalla struttura alla sua incarnazione, dalle condizioni di problemi ai casi di soluzione, dagli elementi differenziali e dalle loro connessioni ideali ai termini attuali e alle relazioni reali diverse che costituiscono ad ogni istante l’attualità del tempo” (Ibidem).

A questo punto il concetto di differenza $\frac{t}{z}$ si biforca in due movimenti:

- i) il primo consiste nella “determinazione del contenuto virtuale dell’Idea”, dove “le varietà dei rapporti si incarnano in specie distinte” (Ivi, p. 333). Deleuze lo denomina *Differenziazione*.
- ii) Il secondo movimento, derivante dal primo, consiste in un nuovo tipo di distinzione, quella cioè in cui il contenuto virtuale dell’Idea si attualizza in specie e parti distinte. Qui si parlerà di *Differenziazione*.

Nessuno dei due processi contempla il negativo, la negazione, la determinazione negativa: il primo processo descrive una positività pura fatta di “elementi di affermazione genetici o produttori” (posti, funzioni, soglie, punti singolari...); il secondo riflette “la produzione di affermazioni generate finite”, cioè quei termini attuali che vanno a occupare posti e posizioni aperti dal primo movimento. Se vi è un negativo apparente, può esservi solo nei termini attuali e nelle relazioni reali poiché “recisi dalla virtualità” (Ivi, p. 334).

Come anticipato, tra le due trattazioni del virtuale si insinuano dei paradossi e delle contraddizioni apparenti. È quindi lo stesso Deleuze a ritenere di dover fornire dei chiarimenti, in quest’ordine: il virtuale non pertiene all’indeterminato, ma alla determinazione della differenza; “il virtuale non si oppone al reale”, ma all’attuale e possiede realtà piena; “il virtuale è parte integrante dell’oggetto reale”; la realtà del virtuale è la struttura e i suoi elementi, rapporti differenziali e punti singolari che non hanno attualità, ma conservano pieno statuto di realtà (Ivi, p. 336).

E ancora, occorre fare attenzione a non confondere il virtuale con il possibile, poiché il possibile si oppone al reale, cosa che invece non interessa la virtualità; inoltre, il possibile è dell’ordine del concetto, il virtuale è dell’ordine della molteplicità nell’Idea. Il processo di realizzazione (dal possibile al reale) si articola per immagine e somiglianza; l’attualizzazione del virtuale avviene per

differenza, divergenza o differenziazione. Rompendo con la rassomiglianza, l'attualizzazione è sempre creazione. Rispetto ai modi di esistenza semiotici, in Deleuze, virtuale e potenziale sono sinonimi; entrambi si attualizzano creando linee di divergenza che corrispondono, ma non rassomigliano al molteplice.

4.2 Differenza e differenziale

C'è un ulteriore aspetto di cui occorre tenere conto nell'evoluzione del pensiero di Deleuze sul virtuale intensivo, un aspetto che viene messo in evidenza da Sarti in occasione della presentazione del libro *Differential Heterogenesis. Mutant forms, sensitive bodies*¹⁶⁰ nel marzo 2023 all'interno degli incontri del gruppo di Quinto Alto.¹⁶¹ Qui Sarti individua uno scarto tra il pensiero di Deleuze in *Differenza e ripetizione* e il pensiero di Deleuze nell'incontro con Guattari in *Mille piani*; questo scarto sta nel fatto che il primo si dipana all'interno di una *filosofia della differenza*, il secondo dentro una *filosofia del differenziale*. Dalla prospettiva di una rappresentazione dinamica, spiega Sarti, la differenza è la “differenza tra un attrattore e un altro attrattore” che si dà in uno spazio esteso. La filosofia del differenziale espressa in *Mille piani* è invece, sempre dal punto di vista delle dinamiche, è una “filosofia della composizione delle forze sul piano intensivo”

Per precisazione, trascriviamo le parole di Sarti: “Il differenziale è un rapporto di differenze che al limite vanno a zero, ma la cosa straordinaria è che benché queste differenze nel loro rapporto vadano a zero, il rapporto non va a zero”. Quello verso cui tende il rapporto tra queste due grandezze differenziali è una “grandezza intensiva”. Poi riprendendo uno dei corsi che Deleuze ha tenuto a Paris 8 negli anni '80 cita questo passaggio:

“Ces quantités intensives s'expriment, se définissent, uniquement, par leur distance à zéro (...). Dès lors, c'est complètement normal que si les essences sont des quantités intensives, elles s'expriment dans des rapports différentiels, puisque la quantité intensive est inséparable d'une définition par rapport à zéro. Et que le rapport différentiel, c'est précisément ça. Tout devient lumineux!” (Sarti et al. 2022, p. 17 n. 2).

Guardando al testo *Mille piani*, il passaggio da una *filosofia della differenza* a una *filosofia del differenziale* ci sembra sia ben testimoniato nella sua ragion d'essere in un punto dell'argomentazione

¹⁶⁰ Sarti, A., Citti, G., & Piotrowski, D. (2022). *Differential Heterogenesis*. In A. Sarti, G. Citti, & D. Piotrowski, *Differential Heterogenesis* (pp. 55–96). Springer International Publishing. https://doi.org/10.1007/978-3-030-97797-9_4

¹⁶¹ Incontri di Quinto Alto (Regista). (2023, marzo 14). A Sarti «*Differential heterogenesis. Mutant forms, sensitive bodies.*» [Film]. <https://www.youtube.com/watch?v=ibUSkIBskjE> [ultimo accesso: 15 ottobre 2025]

in cui si analizza il rapporto tra filogenesi e ontogenesi del vivente: “Non sappiamo più molto bene dove siamo. Ci sono tante cose in gioco in queste repliche pronte. Ci sono tante distinzioni che continuano a proliferare. Ci sono tanti regolamenti di conti, perché l’epistemologia non è innocente” (Deleuze & Guattari 1980, p. 72). Una chiave di lettura di queste distinzioni che si sono moltiplicate in tutti i sensi viene ricercata in un ‘fenomeno decisivo’ nell’ambito dell’evoluzionismo e nel neoevoluzionismo: “il nostro teatro di marionette diventa sempre più nebuloso, cioè collettivo e differenziale. I due fattori che invochiamo con le loro relazioni incerte, per spiegare la diversità sopra uno strato – i gradi di sviluppo o di perfezione e i tipi di forme – subiscono una profonda trasformazione. Secondo una doppia tendenza, i tipi di forme devono comprendersi sempre di più a partire da popolazioni, mute e colonie, collettività o molteplicità; i gradi di sviluppo devono comprendersi in termini di velocità, di tassi, di coefficienti e di rapporti differenziali. Doppio approfondimento. Ecco l’acquisizione fondamentale del darwinismo, ed essa implica un nuovo accoppiamento individui-ambienti sullo strato” (Ivi, p. 72).

Il “teatro di marionette” richiama le dinamiche strutturali dominate dal principio di differenza e controllabili in virtù di parametri esterni allo spazio di fase - o spazio delle variabili di stato interne - nella determinazione di quale attrattore debba o possa prevalere sull’altro. Qui lo spazio di possibilità è chiuso a priori; le forme emergenti saranno delle più disparate nella loro estensione nello spazio e nel tempo, ma lo spazio di possibilità resta vincolato, direbbe Sarti. Nel nuovo scenario che delineano Deleuze e Guattari, avviene invece qualcosa di simile a un rovesciamento che loro descrivono a partire dal rapporto tra embriogenesi e filogenesi: “non è più l’embrione che testimonia di una forma assoluta prestabilita in un ambiente chiuso, è la filogenesi delle popolazioni che dispone di una libertà di forme relative, non essendo nessuna di esse prestabilita in un ambiente aperto” (Ibidem). La ricomposizione cioè si dà sul piano del virtuale intensivo, secondo quella che Sarti chiama “integrazione di un sistema di elementi differenziali” (Sarti et al. 2022, p. 18). Nei termini di Deleuze e Guattari, la diversità tra i tipi di forme non è anteriore all’esistenza stessa delle popolazioni, elementari o molecolari che abitano un ambiente dato, ma le forme sono il risultato di suddivisioni interne all’ambiente, “in quanto la sua molteplicità si dividerà in molteplicità grazie a una differenza di natura”. Al tempo stesso, i gradi di sviluppo anch’essi non sono preesistenti, ma sono equilibri relativi e globali (globali in quanto estesi) e occorre che siano intesi in funzione dei vantaggi che si determinano tra elementi, molteplicità nell’ambiente e variazioni nell’ambiente. E la misurazione dei gradi non segue una scala ascendente di perfezione, ma procede per rapporti e coefficienti differenziali.

Nella traduzione dell'eterogenesi differenziale come idea filosofica del divenire delle forme in una dinamica reale o formalizzazione matematica, Sarti e Citti propongono quindi di sostituire le strutture, gli spazi di possibilità chiusi e predeterminati, con *assemblaggi di spazi di possibilità*. A questo scopo accolgono il superamento della geometria differenziale riemanniana verso le geometrie sub-riemanniane, poiché

“Riemannian manifolds are in some way limited with respect to the kind of heterogeneity they are capable of instantiating, since they are lacking certain crucial features that would enable a properly heterogeneous differential becoming; more specifically, the major limitation in question consists in the fact that the manifold has the same dimension in every point of the space. In this sense, the differential constraints imposed by the metric, although different, result homogeneous in its dimension as the tangent planes have the same dimension point by point. If we consider the classical problem of integration of a curve starting from its tangents, the differential constraints are homogeneous, or more precisely, equiregular” (Ivi, p. 55).

Nel Deleuze di *Differenza e ripetizione*, la *varietà (manifold)* riemanniana era lo strumento di riferimento per la rappresentazione geometrica dell'Idea in quanto molteplicità (Sarti et al. 2022; Deleuze 1968). A differenza dello spazio euclideo, la geometria differenziale di Riemann tentava di superarne l'uniformità e l'isotropia (cioè l'uguale distribuzione interna, poco adatta ad esprimere rapporti differenziali) – ricordiamo per inciso come gli spazi latenti dei modelli di linguaggio contestuali risultino essere fortemente anisotropici.

Localmente, una varietà è una curvatura di un sottoinsieme dello spazio Euclideo o meglio, una deformazione senza tagli, che può essere rappresentata da una *carta* (o mappa locale, Figura 19). Tuttavia, una varietà non può essere mappata su una singola carta locale.

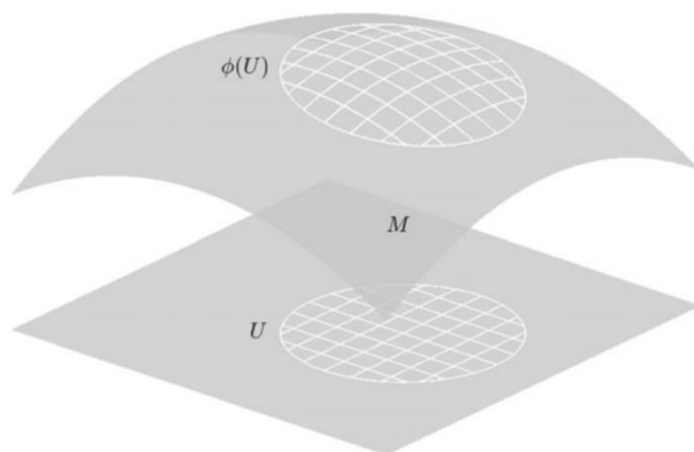


Figura 20: Varietà localmente definita e mappata su una carta singola.¹⁶²

Per definirla sono necessari un insieme di *carte*, chiamate *atlante*. L'esempio tipico è il globo terrestre: se viene mappato su una sola carta, i poli diventano linee. Servono almeno due mappe per coprire una *varietà*, ma possono anche essere di più. E questo è il primo livello di molteplicità: la molteplicità discreta delle mappe locali.

Ora, per comporre due carte sovrapposte in un atlante, per poterle incollare tra loro, è necessario che vi sia un'applicazione liscia, cioè differenziabile: questo significa che se ciascuna delle due carte fornisce due deformazioni diverse della stessa regione della varietà, le due rappresentazioni per poter essere ben *incollate*, devono combaciare senza discontinuità, senza che vi siano “spigoli” o “salti” nel passaggio tra una carta e l'altra. In questo senso, la molteplicità è proprio un modo di ripensare lo spazio non come qualcosa di già dato, ma come una giustapposizione di pezzi (es. carte) che possono essere combinati in modi differenti. Uno spazio eterogeneo appunto. Lo spazio è liscio se la varietà è differenziabile. Laddove è possibile definire una metrica per determinare una distanza in M , allora lo spazio è striato. Sarti e Citti riprendono da *Mille piani* questa distinzione: da un lato una varietà liscia che è struttura locale e continuità, senza una misura, dall'altro uno spazio striato che rappresenta una struttura più rigida in cui è presente una metrica; dunque, qualcosa che ne consente la misurazione (Ivi, p. 40; confronta anche in questo testo p.). Entrambe le varietà riemanniane posseggono eterogeneità al loro interno, ma di natura diversa.

Per introdurre le geometrie sub-riemmaniane, dobbiamo chiarire cosa si intende per “assemblaggio di spazi di possibilità” a cui abbiamo fatto riferimento sopra e che dipendono da quello che viene definito piano di composizione. Facciamo un passo indietro: abbiamo visto che nelle dinamiche morfogenetiche strutturali à la Deleuze di *Differenza e ripetizione* - ma ancora più

¹⁶² In Sarti et al. 2022, p. 33. La funzione che definisce il grafico locale è $\phi: U \rightarrow M$.

rappresentative sarebbero la teoria delle catastrofi di René Thom o le metamorfosi goethiane – il punto era controllare un gruppo dato di singolarità e determinare il passaggio da un bacino di attrattori all'altro per stabilizzare la dinamica, nelle dinamiche post-strutturali o eterogenetiche *à la Deleuze et Guattari*, si tratta di “creare costellazioni di singolarità”, anche e soprattutto aggiungendo campi di forze che possono integrarsi o meno con campi di forze preesistenti.

Così Sarti et al. descrivono questo passaggio: “To allow a mutation of the space of possibilities, it is necessary a transformation of the virtual that generates it, which goes well beyond the parametric variation. As long as we are in the regime of parametric variation we remain within the same space of possibilities, where the mathematical framework proposed by R.Thom is well suitable to take into account for example phase transitions. But to allow the transformation of the phase space and introduce a new one it is necessary to take into account a recomposition of the virtual that is not common to mathematical physics and either in structural morphodynamics” (Ivi, p. 42).

In questa trasformazione del virtuale intervengono due processi concorrenti: da un lato “the specification of adjunct fields on the virtual plane”, dall'altro “the condensation of singularities in its actualization” (Ibidem). Questa idea era già presente nel Deleuze di *Differenza e ripetizione*. L'integrazione dei campi di forze aggiunti è dell'ordine dell'evento, ma soprattutto della sua rarità. La funzione di questi campi aggiunti è quella di completare le condizioni affinché il problema sia risolvibile.

Ci permettiamo, prima di procedere nell'argomentazione matematica, un affondo sul concetto di rarità che, dal punto di vista di una teoria dell'enunciazione (cfr, capitolo 7) è uno snodo importante, soprattutto in relazione all'idea di Enciclopedia come virtuale intensivo. Paolucci (2020) vedremo che riprende da Foucault la caratteristica rarità degli enunciati o “la forma lacunosa e frammentaria del campo enunciativo” (Foucault 1969). La rarità dell'evento in Deleuze assume una diversa configurazione. Questo confronto è stato anche uno dei punti affrontati durante una delle séances del Séminaire “*Dynamiques post-structurelles*” organizzato da Alessandro Sarti e Giuseppe Longo per il Centre d'analyse et de mathématique sociales (CAMS) dell'EHESS. L'incontro dal titolo “*Comment un système peut-il être une hétérogénése?*” - che vedeva come relatori Mattia Galeotti et Francesco Demitry dell'Università di Bologna e che presentava una strutturale corale di interventi - è stato occasione per approfondire quest'idea di rarità e del suo statuto. Come notano Sarti et al. (2022, pp. 123-126), avere un'idea è un “evento raro” (Deleuze 1968), assimilando a questo tipo cognitivo, l'occorrenza dell'insight nella risoluzione di un problema. Qui coesistono i due processi concorrenti cui accennavamo poco sopra: “It is the sublime occasion of Kairos which makes the solution explode like something abrupt, brutal and revolutionary. Hence, the Deleuzian idea that the creative act would

take place as a composition of adjoined differential fields can be found in brain dynamics when different cortical assemblies find a way to join with one another. If we were in the presence of a network, we could refer to this as an act of ‘connection’; as we’ve already considered, however, the wet brain is a differential neuromagma that goes far beyond the connective scheme of the network and contains all the uncertainties and indeterminacies of material concatenations” (Ivi, p. 125). A quest’idea di rarità appartenente al processo di integrazione dei campi aggiuntivi sul piano del virtuale intensivo, sembrerebbe contrapporsi “l’effetto di rarità” che Foucault prende in considerazione a proposito dell’analisi enunciativa e delle formazioni discorsive: “il fatto che, in fin dei conti, solo poche cose possano esser dette, spiega come gli enunciati non siano una trasparenza infinita come l’aria che respiriamo; ma cose che si trasmettono e si conservano, che hanno un valore, e che si cerca di far proprie; che si ripetono, si riproducono e si trasformano; per le quali si predispongono dei circuiti prestabiliti e alle quali si dà statuto nell’istituzione; cose che vengono reduplicate non soltanto dalla copia o traduzione, ma dall’esegesi, dal commento e dalla proliferazione interna del senso”¹⁶³. Nell’*Ordine del discorso* (Foucault 1970) mette in guardia dal guardare alle rarefazioni del discorso come effetti di una censura, un non-detto originario da portare alla luce. La natura produttiva della legge di rarità sta nel suo valore metodologico: dispersione, frammentarietà e discontinuità definiscono il campo positivo dell’analisi degli enunciati, insieme ad altri due concetti chiave, quello di exteriorità (come nel valore saussuriano) e quello di accumulazione (costituzione di archivi). Sottolineiamo, infine, che quest’aspetto epistemologico riferito alla finitezza strutturale degli enunciati in Foucault è condizione della loro regolarità (e non regola).¹⁶⁴

Ora, la contrapposizione che per noi sussiste è la seguente: la rarità dell’evento in Deleuze pertiene direttamente alla virtualità che è il “luogo” dove i campi aggiunti si integrano e si ri-compongono. Ciò che è raro su questo piano è proprio il fatto che si dia integrazione e composizione tra campi di forze. In Foucault la lacunosità e la rarità interessano gli enunciati già realizzati, cioè il fatto che non tutto possa essere contemporaneamente detto, proprio perché la virtualità è troppo piena e deve essere ritagliata per “aggiunta di sottrazioni”. Nel virtuale dell’enunciazione il senso non è mai mancante, vi è al contrario un più di senso (Paolucci 2021; 2020). Da qui, si apre l’interrogativo sul modo di esistenza della rarità.

È interessante allora notare che nel suo testo su Foucault, Deleuze (1986) dichiara che la rarità degli enunciati è una realtà di fatto e di diritto, cercando forse in questo modo di sciogliere la

¹⁶³ Foucault, M. 1969, *L’archéologie du savoir*, Éditions Gallimard; tr. it. *L’archeologia del sapere. Una metodologia per la storia della cultura*, BUR Rizzoli, 2009, p. 111.

¹⁶⁴ Su questo si veda ancora Paolucci (2020).

contraddizione. Si parla di diritto, poiché la rarità è legge ed effetto della produzione di enunciati, in ciò distinguendoli dalle proposizioni e dalle frasi. Le proposizioni sono soggette a un processo di realizzazione¹⁶⁵, cioè di passaggio dal possibile al reale, che ne consente un pullulare continuo e un'espressione "le une 'sulle' altre conformemente alla distinzione di tipi" (Deleuze 1986, p. 9). La rarità di fatto concerne, invece, "ciò che è realmente detto, cioè "una frase ne nega delle altre, ne impedisce delle altre, contraddice o respinge altre frasi; benché ogni frase sia gravida anche di tutto ciò che non dice, di un contenuto virtuale o latente che ne moltiplica il senso e si offre all'interpretazione dando luogo a un «discorso nascosto», vera e propria ricchezza di diritto" (Ivi, p. 10). Egli parla di una dialettica delle frasi e di un'astrazione per tipologia delle proposizioni: la prima come possibilità di opporre una frase a un'altra (per risolvere una contraddizione), la seconda come possibilità di stratificare proposizioni l'una sull'altra.

Nel caso degli enunciati, invece, Deleuze parla di un principio di parsimonia o anche di deficit; dunque, nella sua lettura di Foucault gli enunciati sono inseparabili da uno spazio di rarità perché non c'è "né un possibile né un virtuale, ma tutto è reale e ogni realtà è manifesta" (Ibidem). Lo spazio di rarità è soggetto anch'esso a delle opposizioni e gerarchizzazioni interne, ma "in ragione di una distanza positiva misurabile nello spazio di rarità". Tali confronti avvengono in virtù di una "diagonale mobile" che consente di selezionare uno stesso insieme attraversando livelli differenti oppure insiemi differenti appartenenti al medesimo livello. Ed è in virtù della rarità che non solo si dicano poche cose, ma che "poche cose *possano* essere dette" (corsivo nel testo). Tuttavia, la produzione di enunciati non presenta alcuna originalità, ma "Ciò che conta è la regolarità dell'enunciato: non una media, ma una curva. *L'enunciato infatti non deve essere confuso con l'emissione di singolarità che presuppone*, ma con l'andamento della curva che passa in vicinanza di queste singolarità, e più in generale con le regole del campo in cui si distribuiscono e si riproducono" (Ivi, p. 11 corsivo nostro). Infine, non vi è un'origine dell'enunciato, nessun Io che dice Io per la prima volta, ma in ciascuno enunciato ci sono dei "posti" di soggetto che possono essere occupati da individui diversi (Cfr. Paolucci, 2020); "l'enunciato è l'oggetto di un cumulo attraverso cui si conserva, si trasmette o si ripete" e tale cumulo proprio è un effetto della rarità.¹⁶⁶

¹⁶⁵ Ricordiamo che per Deleuze il virtuale/potenziale non coincide con il possibile: il virtuale si attualizza, il possibile si realizza.

¹⁶⁶ In altra sede, varrebbe la pena rileggere quest'intersezione alla luce del concetto di virtuale, trattazione sistematica assente in Foucault, ma che ha visto linee interpretative differenti nel contesto francese (Cfr. Sforzini, A. (2014). *Michel Foucault entre histoire et fiction*. In M. Panter, P. Mounier, M. Martinat, & M. Devigne (eds.), *Imagination et histoire: enjeux contemporains*. Presses universitaires de Rennes. <https://doi.org/10.4000/books.pur.49586>) e anglofono (cfr Penfield, C. (2023). Foucault's Virtual Force Ontology. *The Journal of Speculative Philosophy* 37(3), 447-457.

Ci sembra molto chiara la distinzione qui tra enunciazione, enunciato e atto di enunciazione, dove l'enunciazione istituisce le condizioni di possibilità sia del verificarsi dell'atto che dell'esistenza dell'enunciato, nel senso “delle condizioni di singolarità che sono presupposte”. L'atto rappresenta l'evento e l'enunciato manifesto è “forma lacunosa e frammentaria”.

Riprendiamo ora la nostra argomentazione matematica sulla ricomposizione del virtuale. L'evento può anch'esso essere definito strutturale, quando il flusso di attualizzazione si determina nel passaggio tra due bacini stabili, in presenza di un dispositivo di controllo. Si tratta di un passaggio da una struttura ad un'altra, che conserva “the deep functioning of the structure as a device for categorization. René Thom has shown clearly that the emergence of structures requires a very particular kind of singularities defining stable basins of attraction”. (Sarti et al. 2022, p. 43).

Un'altra possibile attualizzazione della composizione dei campi aggiunti deriva dalle singolarità generali di Poincaré che differiscono dalla dinamica di potenziale gradiente, alla base della morfodinamica strutturale, poiché quest'ultima evolve sempre in “discesa” verso un minimo di potenziale che garantisce la stabilizzazione morfologica. Nel caso della singolarità di Poincaré si può parlare di un insieme di condizioni critiche in cui il comportamento del sistema non è riconducibile a un gradiente, ma apre lo spazio a dinamiche non conservative, ricorsive, caotiche, non integrabili.

$$\frac{d}{dt}x(t) = -\nabla V(x, p), \quad \frac{d}{dt}x(t) = Ax(t),$$

Figura 21: Evoluzioni di due sistemi.¹⁶⁷

<https://muse.jhu.edu/article/904128>). Con ciò intendiamo dire che sembra non esserci un pensiero del virtuale in Foucault, con poche eccezioni tra le quali ricordiamo le eterotopie, ossia quei “luoghi assolutamente altri” che per Foucault esistono “in ogni cultura come in ogni civiltà, dei luoghi reali, dei luoghi effettivi, dei luoghi che appaiono delineati nell'istituzione stessa della società, e che costituiscono una sorta di contro-luoghi, specie di utopie effettivamente realizzate nelle quali i luoghi reali, tutti gli altri luoghi reali che si trovano all'interno della cultura vengono al contempo rappresentati, contestati e sovvertiti; una sorta di luoghi che si trovano al di fuori di ogni luogo, per quanto possano essere effettivamente localizzabili”. (Foucault, M. (2010). *Eterotopia*. Mimesis. <https://web.unica.it/unica/protected/417979/0/def/ref/MAT417974/>). I contro-luoghi verrebbero a rappresentare, a livello di organizzazione spaziale, quella tensione tra virtuale e reale che abbiamo descritto nel paragrafo precedente.

¹⁶⁷ A sinistra: evoluzione di un sistema gradiente, dove $x(t)$ è la variabile di stato, $d/dt x(t)$ è la sua variazione nel tempo, $\nabla V(x)$ è il gradiente del potenziale $V(x)$ e p è lo spazio dei parametri di controllo. A destra: evoluzione di un sistema dinamico lineare, dove x è lo stato del sistema, $d/dt x(t)$ è l'evoluzione del sistema nel tempo e A è la matrice di evoluzione.

Nel caso delle dinamiche post-strutturali, invece, la composizione dei campi aggiunti che formano assemblaggi sostituisce il dispositivo di controllo dello spazio di fase presente nelle dinamiche strutturali. Le relazioni interne ed esterne non sono oppositive, come in una struttura, ma si moltiplicano nella forma di connessioni, congiunzioni e alleanze. Dicono Sarti et al.: “An assemblage is thus an a-structural genetic element for any dynamic individuation, the composition and actualization of which give rise to differential heterogenesis”. (Ivi, p. 44). E più avanti: “The term assemblage in *A thousand plateaus* is declined both from a geometric point of view, in terms of *rhizome*, and from a dynamic perspective, in terms of *machine* or *machinic assemblage*, that is, the heterogeneous becoming of forms from a primordial chaotic soup” (Ivi, p. 45). Approfondiremo il concetto di assemblaggio nel prossimo capitolo 5.0; in relazione al rizoma, abbiamo già visto il ruolo che esso aveva nella rappresentazione dell’Enciclopedia per lo stesso Eco. Ci preme sottolineare che tanto il rizoma, quanto l’Enciclopedia presentano una dimensionalità non predeterminata, ma che si modifica sulla base delle relazioni tra campi e relazioni interne ed esterne. In questo senso non sono strutture, ma possono coesistere con strutture.

Poste le premesse teoriche che compongono la nostra idea di assimilare l’Enciclopedia echiana al virtuale intensivo deleuziano, proviamo a trarre delle conseguenze, non tanto nella forma di conclusioni, quanto di condizioni che ci permettano di argomentare l’analogia tra queste configurazioni discorsive del virtuale e il modello di mondo che le macchine dotato di linguaggio hanno appreso.

5.0 Dell'IA in sé tra strutture e assemblaggi

Introduciamo a questo punto alcune delle teorie del linguaggio che fanno da sfondo alle analisi interpretative circa la capacità dei Large Language Models di generare e interpretare enunciati basati sul linguaggio naturale. Tra le varie ipotesi ci ha colpite, in particolare, la provocazione di Ambridge e Blything (2024)¹⁶⁸ a considerare i LLM “the leading current theories of how speakers learn and represent these restrictions [il riferimento sono i giudizi di accettabilità grammaticale]. Of course, they are not perfect theories – far from it – but they’re better theories than any others that have been proposed” (Ivi, p. 34), in quanto in grado di replicare le intuizioni linguistiche dei parlanti umani. Più nello specifico per la loro capacità di “learning and representing verbs’ argument structure privileges” (Ivi, p. 33), definite come valenze verbali (transitive e/o intransitive). Questa posizione fa eco a quelle sostenute da Piantadosi (2022¹⁶⁹; 2024¹⁷⁰), il quale, oltre ad aver dichiarato il sovvertimento e superamento del quadro teorico della linguistica generativa ad opera dei “moderni modelli di linguaggio”, in accordo a una certa teoria della mente, sostiene che i LLM possono accedere al significato non per via referenziale (vedi 0.1), ma attraverso il ruolo concettuale dei lessemi che è definito dalle relazioni tra gli stati di rappresentazione interni al modello. Nello stesso articolo del 2024, Piantadosi discute lo status dei Large Language Models come teorie scientifiche in questi termini:

“Their parameters come to embody a theory of language, including representations of latent state through a sentence and a discourse. The exact same logic of tuning parameters to formalize and then compare theories is found in other sciences, like modeling hurricanes or pandemics: any set of assumptions will generate a distribution of predictions and the assumptions are adjusted to make the best predictions possible. In this way, a learning mechanism configures the model itself in the space of theories in order to satisfy the desired objective function. For hurricanes or pandemics, this is as

¹⁶⁸ Ambridge, B. & Blything, L. (2024). Large language models are better than theoretical linguists at theoretical linguistics. *Theoretical Linguistics*, 50(1-2), 33-48. <https://doi-org.ezproxy.unibo.it/10.1515/tl-2024-2002>. Loro stessi definiscono il titolo come un “clickbait”.

¹⁶⁹ Piantadosi, S. T., & Hill, F. (2022). *Meaning without reference in large language models* (No. arXiv:2208.02957). arXiv. <http://arxiv.org/abs/2208.02957>

¹⁷⁰ Piantadosi, S. (2024). *Modern language models refute Chomsky’s approach to language*. LingBuzz. <https://lingbuzz.net/lingbuzz/007180> La tesi centrale è che la rottura operata dai LLM non è parziale, ma investe le fondamenta del paradigma generativo chomskiano, incluse le sue principali implicazioni: “none of the principles and innate biases that those who work in that tradition have long claimed necessary needed to be built into these models (e.g. binding principles, binary branching, island constraints, empty category principle, etc.). Moreover, these models were created without incorporating any of Chomsky’s key methodological claims, like ensuring the models properly consider competence vs. performance, respect “minimality” or “perfection,” and avoid relying on the statistical patterns of unanalyzed data” (Ivi, p. 15).

rigorous as science gets; for sequences of words, everyone seems to lose their mind. (Piantadosi 2024, p. 9).

Nel rispondere alla critica chomskiana dello “Anything goes” da cui dedurre che questi modelli sono una teoria linguistica, Piantadosi argomenta per esclusione, cioè ogni modello è capace di agire efficacemente in determinati task e non in altri; dunque, il principio dello “Anything goes” è parzialmente rispettato. Tuttavia Piantadosi individua uno dei limiti nel funzionamento dei modelli che da un punto di vista fattuale ci sembra particolarmente rilevante: al di là della performance (più o meno efficace nel reasoning, rispetto al language modelling), c’è un tema centrale che ha grosse ricadute in termini di impatti di questi sistemi, ossia il fatto che siano addestrati “on truly titanic datasets compared to children, by a factor of at least a few thousand (see Warstadt & Bowman 2022 for a comprehensive review of models in language acquisition)” (Ivi, p. 14).¹⁷¹

Di posizione contraria a quest’idea che “the model is the theory” è Müller (2025)¹⁷² il quale negando ai modelli lo statuto di teorie del linguaggio, li descrive piuttosto come strumenti modellistici privi di potere esplicativo: “A theory contains descriptive and explanatory statements about some part of reality. It contains laws about the domain that is described by the theory. A model is an abstract representation of the relevant part of the reality under consideration. A theory can be used to build such models. A theory should use primitives that are appropriate for a certain domain and it should contain statements about these primitives. LLMs are neuronal nets that have been organized and trained in a certain way. Nodes of such nets can be examined and we can even find certain information in them (Zhang & Bowman 2018; Manning et al. 2020), but this information is not a theory. The net may reflect grammatical structures and reject impossible ones, but it does not tell us why this is the case.” (Müller 2025: 2).

¹⁷¹ Su questo ha si veda anche l’intervento di Capone, L. (2024) al XXIX Congresso della Società di Filosofia del linguaggio “Filosofia dei linguaggi e paradigmi psicologici” dal titolo: “*Un’analisi psicolinguistica delle strategie di apprendimento dei Language Model*”, in pubblicazione in Capone, L. (Forthcoming). Child Development and Language Model Training: Comparing Psychological Theories of Learning. *RIFL SFL – Filosofia dei linguaggi e paradigmi psicologici*. Per una descrizione delle più recenti sperimentazioni di addestramento di modelli italiani nativi di baby LMs si vedano: Capone, L., Suozzi, A., Lebani, G. E., & Lenci, A. (2025). *BAMBI Goes to School: Evaluating Italian BabyLMs with Invalsi-ITA*. CLiC-it 2025: Eleventh Italian Conference on Computational Linguistics, Cagliari.; Suozzi, A., Capone, L., Lebani, G. E., & Lenci, A. (2025). *BAMBI: Developing Baby Language Models for Italian* (No. arXiv:2503.09481). arXiv. <https://doi.org/10.48550/arXiv.2503.09481>

¹⁷² Müller, S. (2025). Large Language Models: The best linguistic theory, a wrong linguistic theory, or no theory at all? *Journal of the Linguistic Society of Germany*, 44(1). <https://doi.org/10.18148/ZS/2025-2001>

Ciò che queste posizioni individuano è una distinzione ontologica ed epistemologica tra modello e teoria, che appare dunque fondamentale per orientare correttamente la riflessione semiotica e linguistica sui LLM. Dalla nostra prospettiva, che abbiamo visto non mirare direttamente a definire uno standard di funzionamento unificato del significato (Cfr. Violi 1997), ma a fornire una descrizione efficace della competenza semiotica di questi modelli, si tratta di dare conto degli *effetti di senso* che l'IA generativa intrattiene con i suoi testi e che noi intratteniamo con l'IA generativa in quanto testo e pratica semiotica. Così Pezzini (2007)¹⁷³ definisce l'effetto di senso: “generalmente si parla di ‘effetto di senso’ per indicare il risultato dell’interazione fra lettore-interprete e testo, che può diventare il punto di partenza per andare a scoprire quali articolazioni e organizzazioni semio-linguistiche sono in grado di produrlo” (Ivi, p. 159). Ci piace anche richiamare in questa sede la posta in gioco di questi effetti di senso, come formulata da Vetere (2025): con l’ingresso in scena delle macchine parlanti, al di là dei problemi rimasti irrisolti delle teorie e filosofie del linguaggio, si tratta di “come vivere in un ambiente popolato da nuovi s-oggetti che fanno un loro peculiare esercizio della nostra più umana facoltà” (Ivi, posizione 526).

Nel resto del suo libro, Vetere intreccia teorie del linguaggio e sviluppo dell’intelligenza artificiale, a partire dagli anni '50 – gli anni delle profezie di Alan Turing (1950)¹⁷⁴ e la conferenza di Dartmouth (1955)¹⁷⁵ – identificando di fondo nel volgersi del dominio informatico al linguaggio una seconda “svolta linguistica”:

“Agli inizi del Novecento la filosofia diventa in larga misura una filosofia del linguaggio, non solo perché sorge il dubbio che i suoi problemi siano in fondo ‘grattacapi linguistici’, ma anche perché si inizia a riflettere sulla circostanza che il discorso filosofico è fatto in fin dei conti di parole. Allo stesso tempo, la linguistica esce dalla *comfort zone* della filologia e della glottologia comparativa,

¹⁷³ Pezzini, I. (2007). *Il testo galeotto: La lettura come pratica efficace*. Meltemi.

¹⁷⁴ Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind, New Series*, 59(236), 433–460. Con queste parole si esprimeva Turing a proposito del futuro delle macchine capaci di pensare: “I believe that in about fifty years' time it will be possible, to programme computers, with a storage capacity of about 10⁹, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning. The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. I believe further that no useful purpose is served by concealing these beliefs” (Ivi, p.442). Ci sembra che il margine di errore dell’ordine di un paio di decenni sia fondamentalmente trascurabile.

¹⁷⁵ McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*: August 31, 1955. *AI Mag.*, 27(4), 12–14. <https://doi.org/10.1609/aimag.v27i4.1904>

cioè del censimento e della classificazione dei "piccoli fatti veri" del linguaggio, addentrandosi nella speculazione sulla lingua come sistema simbolico che sottende alla vita individuale e sociale. Dunque avviene che filosofi e linguisti si trovino a convergere fundamentalmente nella stessa ricerca. Cento anni dopo, cioè oggi, anche l'informatica svolta verso il linguaggio, in quanto, sotto le insegne della cosiddetta *intelligenza artificiale generativa*, prende la forma di una tecnologia il cui paradigma è quello della produzione di strutture simboliche lineari con metodi mutuati dall'ingegneria del linguaggio naturale. Tra le due 'svolte linguistiche', che si susseguono a distanza di un secolo, ci sono connessioni importanti. Queste aiutano a capire la presenza già attuale degli automi linguistici nella nostra società e il posto che questi potranno occupare nel prossimo futuro" (Ivi, posizione 157).

All'interno dell'arco temporale preso in considerazione da Vetere, il nostro sguardo semiotico coglie un punto di snodo che non appartiene tanto alla storia delle idee e delle teorie, quanto ad un certo modo di concepire la dimensione morfologica e spaziale del linguaggio, della cognizione e del pensiero. Come in altra occasione ci è già capitato di affermare, parliamo di forme, forse rapite da quell'affermazione che Paul Valéry fa pronunciare a Socrate nel dialogo "Eupalino o dell'Architettura". Così Socrate diceva a Fedro: "per noi Greci, tutto è forma. Noi, non consideriamo che rapporti e, come immersi in una luce limpida, simili a Orfeo, col mezzo della parola costruiamo templi di saggezza e di scienza che possono bastare a tutti gli esseri ragionevoli" (Valery, P. 1921). Provocatoriamente ci domandiamo in sostanza cosa non abbia una forma o cosa non sia, in qualche modo, interpretabile per il tramite di forme o rapporti.

5.1 Lo schematismo posizionale morfodinamico e la sintassi di attrattori

Nella traiettoria delle teorie del linguaggio tracciata da Vetere, scegliamo di soffermarci sul modello delle *dipendenze sintattiche* – come le definisce lo stesso Vetere - teorizzate dal linguista francese Lucien Tesnière. Lo facciamo per una ragione molto precisa, che ha a che fare con i passaggi di dimensionalità tra "il parlare" e "il comprendere". In *Éléments de syntaxe structurelle*, Tesnière definiva il parlare una lingua come "transformer l'ordre structurel en ordre linéaire" e viceversa il comprendere una lingua come "transformer l'ordre linéaire en ordre structurel" (Tesnière, L. 1959, p. 19).

Questi passaggi sono guidati dal trasferimento delle connessioni proprie dell'ordine strutturale alla sequenza dell'ordine lineare, secondo un principio di "voisonage immédiat sure la chaîne parlée" (Tesnière 1959, 20).

Tra i due ordini vi è un'antinomia di fondo, determinata dalla possibilità per un termine dell'ordine strutturale di avere più di due connessioni, oltre a quella unica di ordine superiore, là dove una parola

nella sequenza lineare della frase non può che connettersi al suo antecedente e al suo conseguente. Dice Tesnière “En d’autres termes, tout nœud structural est susceptible de créer des bifurcations, trifurcations, etc., incompatibles avec l’ordre linéaire.” (Ivi, 21).

L’antinomia scaturisce dunque da un problema di passaggi tra dimensionalità: l’ordine strutturale è a più dimensioni (rappresentate bidimensionalmente in quelli che Tesnière chiama *stemmi*, che altro non sono se non rapporti di dipendenza gerarchica tra le parole che compongono la frase). Al contrario, l’ordine lineare è unidimensionale. La risoluzione dell’antinomia per il linguista francese è la *conditio sine qua non* della *parole*.

All’intima importanza delle connessioni per la formalizzazione delle strutture profonde diede attenzione Jean Petitot, il quale in un articolo dal titolo “Approche morphodynamique de l’iconicité des stemmas. Des connexions tesnières aux images-schémas des grammaires cognitives” (Petitot 1995)¹⁷⁶ così scriveva: “ces connexions sont des incorporels qui ne sont indiqués par rien (1.4), qui ne possèdent aucun marquant (16. 12) et ne peuvent être saisis que par «l’esprit»”. Tali connessioni non posseggono una relazione di tipo logico, aprendo il ragionamento di Petitot al problema dello statuto eidetico di questa relazione. Dal canto suo, Tesnière dichiarava esplicitamente che un’economia dell’ordine sintattico era comprensibile solo a chi avesse accesso alla totalità dei suoi elementi costitutivi alla ricerca della legge interna della loro disposizione. Un inciso anticipatorio: ciò eccede la competenza umana, evidentemente non quella macchinica.

Per il soggetto parlante, l’attività sul piano strutturale non può che essere analizzata se non attraverso un ritorno introspettivo sull’attività stessa.¹⁷⁷ E ciò induce Tesnière ad appellarsi al metodo

¹⁷⁶ Petitot, J. (1995). *Approche morphodynamique de l’iconicité des stemmas. Des connexions tesnières aux images-schémas des grammaires cognitives*, in Richard-Zappella, J. (1995). *Lucien Tesnière aujourd’hui: Actes du colloque international CNRS URA 1164 [S.U.D.L.A.] - Université de Rouen*, 16, 17, 18 Novembre 1992. Peeters Publishers, pp. 105-112.

¹⁷⁷ Cfr. Paolucci, C. (2006). Lucien Tesnière autore della logica dei relativi. Su alcune insospettite corrispondenze tra Peirce e lo strutturalismo. E/C, 2006, 1-16. In questo contributo, Paolucci propone una lettura *à la Borges* del rapporto tra la fondazione semiotica peirciana e la fondazione semiotica strutturalista. Il concatenamento che dà origine al collassamento delle due anime della semiotica viene identificato nel rapporto tra la logica dei relativi di Peirce e la sintassi strutturale e attanziale di Tesnière. Questo lavoro costituisce per noi una legittimazione teorica di una certa idea di traduzione semiotica del passaggio dal linguaggio naturale all’embedding posizionale e semantico nei LLM. Ci preme però sottolineare un aspetto relativo al metodo introspettivo che Tesnière dichiara qui come necessario: Paolucci riferisce come Peirce ponesse “a fondamento della logica una topologia semiotica fondata costitutivamente sulla dimostrazione delle quattro incapacità e sulla deduzione di alcune sue conseguenze: dal momento che non abbiamo alcun tipo di capacità di intuizione né di introspezione, alcun concetto dell’assolutamente inconoscibile e nessuna capacità di pensare senza segni, occorre allora dividere il flusso continuo del pensiero ed individuare al suo interno delle posizioni che abbiano tra loro un rapporto che sia costitutivamente semiotico e triadico” (Ivi, p. 5).

introspettivo, consapevole del pericolo di carenza di oggettività che esso comporta, nella misura in cui poggia la sua ragion d'essere sull'intuizione come esperienza interna al soggetto parlante. Così descrive la tensione tra oggettivo e soggettivo nello studio della sintassi:

"§4 La méthode introspective se verra reprocher son caractère subjectif, parce qu'elle fait appel à l'intuition.

§5 Ici le grief est encore plus contestable. Assurément, la méthode introspective fait appel à l'intuition. Mais elle fait aussi appel à l'expérience interne. Elle est à ce titre une méthode expérimentale et par conséquent objective.

§6 La méthode introspective est également objective parce qu'elle porte sur des faits. Certes, ces faits sont abstraits. Mais il n'existe pas que des faits concrets. On peut même dire qu'en syntaxe tous les faits sont abstraits, puisque après tout ce n'est jamais que leur marquant¹⁷⁸ qui est concret. Or il n'y a guère d'autre méthode pour atteindre les faits syntaxiques purement abstraits que celle de l'analyse interne" (Tesnière 1959, 38).

Lo stesso Petitot rilevava che le connessioni sono relazioni tanto ideali e astratte, quanto realizzate concretamente all'interno di un soggetto di performance, che "Tesnière disait «inconscientes», on

La definizione di introspezione proposta da Peirce è la seguente: "By introspection, I mean a direct perception of the internal world, but not necessarily a perception of it *as* internal. Nor do I mean to limit the signification of the word to intuition but would extend it to any knowledge of the internal world not derived from external observation" in Peirce, C. S. (1868). Questions Concerning Certain Faculties Claimed for Man. *The Journal of Speculative Philosophy*, 2(2), 103–114.

¹⁷⁸ Il valore di concretezza che caratterizza i fatti della sintassi per Tesnière risiede nella loro marca morfologica (le *marquant morphologique*), ossia la relazione tra ciò che espresso rispetto a ciò che è da esprimersi. Vediamo meglio: nella determinazione del primato tra *exprimende* (il pensiero e il suo schema strutturale e lineare sul piano linguistico) /*exprimé* (vestito fonetico che dona forma sensibile all'*exprimende*) – ossia l'elemento antecedente nel rapporto tra sintassi e morfologia –, la marca costituisce l'inverso del senso. Il senso o la significazione si dà nel rapporto della forma espressa (*exprimé*) a ciò che è da esprimersi (*exprimende*). Dunque, sembrerebbe che il senso debba essere ricostruito a partire dalla forma espressa, il vestito fonetico o la forma sensibile, da cui deriverebbe il primato della morfologia sulla sintassi. Tuttavia, per Tesnière tale primato non può essere riconosciuto, poiché da un punto di vista logico la sintassi precede la morfologia: "Lorsque nous parlons, notre intention n'est pas de trouver après coup un sens à une suite de phonèmes qui lui préexistant, mais bien de donner une forme sensible aisément transmissible à une pensée qui lui préexiste et en est la seule raison d'être" (Tesnière 1959, p. 36). Notiamo una circolarità che ci è familiare in rapporto all'analisi della relazione tra espressione e contenuto che abbiamo fatto in 0.2.4 seguendo Paolucci (2010), a cui rimandiamo con particolare riferimento al capitolo 4: "Semantica e semiotica, espressione e contenuto".

dirait maintenant, dans le cadre d'un cognitivisme computationaliste, «implémentées»” (Petitot 1995, p. 107)

Ma la natura di queste connessioni invisibili e non marcate (a marca zero) per la linea interpretativa Petitot-Paolucci non devono essere assimilate alle relazioni logiche (dalla logica di Port Royal a Frege e Russel), alla correlazione noesi/noema nella fenomenologia husserliana oppure al funzionalismo rappresentazionalista e computazionalista. Dice Petitot:

“Pour lui [Tesnière], la connexion est, on le sait, un principe organique et vital d'organisation (1. 8). Les connexions correspondent à *l'innere Sprachform* de Humboldt. Elles déterminent la production des phrases. Avec les stemma qui les organisent hiérarchiquement, elles «expriment l'activité parlante que l'on a opposée sous le nom de parole au résultat de cette activité (...) qu'est la langue» (3. 11). Je n'ai pas besoin de rappeler ici la façon dont Tesnière reprend ce qu'il appelle «l'intuition de génie» de Guillaume de Humboldt «qui appelait cette opposition de deux mots grecs pleins de sens profond, *l'ergon* (langue) et *l'energeia* (parole)» (Ibidem).

L'associazione tra *energeia* humboldtiana e parole *saussurriana* sancisce per Petitot il primato della performance come principio di produzione, di contro al formalismo della competenza (distinzione cara a Chomsky). Cioè si tratta di individuare una soluzione a un problema che è simile a quello della morfogenesi e che ha a che fare con un quadro di natura cognitivista e meno linguistica. Si potrebbe definire un problema di costruzione del percetto e delle strutture concettuali.

E in risposta a una possibile lettura delle rappresentazioni grafiche stemmatiche di Tesnière come precorritrici degli alberi sintagmatici di Chomsky, Petitot rileva come vi sia solo una parziale affinità. Le connessioni costitutive delle strutture nei costituenti sono delle relazioni formali che generalizzano delle relazioni logiche sulle quali è possibile operare un calcolo simbolico iterativo e trasformatore. Tali descrizioni formali non sono tuttavia in grado di rendere ragione dell'*energeia* linguistica.

Difatti ritornando alla concezione humboldtiana del linguaggio da un lato come “lavoro dello spirito” dall'altro come “organo formativo del pensiero”, è chiaro lo scarto che separa il prodotto (*ergon*) dall'attività (*energeia*). Diremmo noi, il piano dell'enunciazione da quello dell'enunciato.

Descriviamo di seguito gli elementi fondamentali della sintassi strutturale - rivista attraverso una naturalizzazione in chiave di topologia morfodinamica - così come vengono argomentati da Petitot e,

dove ci sembra necessario, approfondiamo con rimandi anche alla formalizzazione matematica proposta dagli autori (a cui Petitot faceva riferimento sin da *Morfogenesi del senso*¹⁷⁹):

1. L'iconicità sintattica e la configurazione posizionale sono legate a una concezione topologica degli stemmi e della valenza. Sarebbe a dire che l'unica astrazione possibile alternativa all'idealità logica è l'idealità spaziale, da Petitot ben distinta dallo spazio fisico. Così, ispirandosi a René Thom, le connessioni di Tesnière sono connessioni posizionali, cioè posizioni occupate da attanti. E i sistemi di connessioni sono delle configurazioni posizionali, ossia delle *Gestalten* posizionali, al cui interno si manifestano trasformazioni temporali (aspettualizzazioni), eventi lessicalizzati dai verbi e rappresentanti degli stati critici che danno luogo all'evoluzione temporale di un processo, tipi di interazione (categorizzazioni) corrispondenti a dei ruoli semantici. Questo approccio topologico, dice Petitot, non è stato pienamente compreso sino allo sviluppo delle grammatiche cognitive e dei lavori sulla schematicità e l'iconicità sintattica di Langacker, Talmy, e Lakoff. In particolare, Langacker e Talmy hanno mostrato che:

“en tant qu'images-schémes catégorisant les états de choses du monde, les structures syntaxiques expriment des contenus conceptuels. D'où l'inséparabilité de la syntaxe et de la sémantique. Mais ces contenus sont évidemment très particuliers. Ce sont des modes d'organisation et de construction des contenus sémantiques au sens standard qui reposent en grande partie, au-delà de toute référence proprement spatio-temporelle, sur un schématisme topologique (Gestaltiste).” (Petitot 1995, 111).

Ciò che evidentemente non può essere ricompresa in questa lucida argomentazione è la dimensione pragmatica della produzione di enunciati, che abbiamo visto essere inscindibile in Eco, che attiene direttamente alla dimensione dell'uso. La svolta testualista non si era ancora compiuta quando Tesnière era in vita (ricordiamo che muore nel 1954 e gli Elementi vengono pubblicati postumi). Per non parlare della semantica del corpus.

2. I potenziali generatori delle connessioni spaziali e i modelli morfodinamici, al cui riguardo Petitot dichiara il contributo fondamentale apportato da Thom (1977)¹⁸⁰ con l'idea dei minimi di potenziale come principio generatore o genetico. Sono i minimi di potenziale a rappresentare le posizioni all'interno di una configurazione posizionale che viene controllata attraverso un principio dinamico produttore, il potenziale generatore appunto. Stiamo, in sostanza, collegando quanto

¹⁷⁹ Petitot, J. (1985). *Morphogenèse du Sens. Pour un schématisme de la structure*. Presses universitaires de France.

¹⁸⁰ Thom, R. (1977). *Stabilité structurelle et morphogénèse. Essai d'une théorie générale des modèles* (II). InterEditions; tr. it. *Stabilità strutturale e morfogenesi. Saggio di una teoria generale dei modelli*. Ghibli (2023).

descritto nel paragrafo precedente a proposito delle dinamiche strutturali con il funzionamento delle dipendenze sintattiche di Tesnière. Vediamo in dettaglio cosa dice Thom (1977, 64 e seguenti):

Sia dato un aperto W dello spazio-tempo R^4 e un campo di dinamiche locali (M, X) , dove M è la varietà-fibra, compatta, lo spazio degli stati interni e X è il campo di vettori associato a ciascun punto x in W che dipende differenzialmente dalle coordinate di x . L'ipotesi di Thom è che il campo $X(x)$ è un campo di gradiente su M :

$$X(m; x) = - \text{grad } V(m; x)$$

dove $V(m; x)$ è una funzione potenziale su M dipendente differenzialmente da x ; il gradiente di V è definito in rapporto a una metrica riemanniana fissata su M . In un punto x di W , i regimi locali stabili sono definiti dagli attrattori strutturalmente stabili del campo X , attraverso i minimi di potenziale $V(m; x)$. In un aperto W vi sono più minimi di potenziale, ma solo uno può governare in un punto x . Come scegliere il regime guida tra tutti i regimi stabili teoricamente possibili? La risposta di Thom risiede nell'adozione della convenzione di Maxwell.

“Se più attrattori c_i , tutti stabili, sono in competizione in un x su W , l'attrattore c_i che prevale è quello di potenziale minimo:

$$V(c_i) < V(c_j) \quad j \neq i$$

L'avevamo già visto a proposito del confronto tra dinamiche strutturali o dinamiche di gradiente e le singolarità generali Poincaré. Poco interessato alla validità quantitativa, per Thom il punto di interesse di questa convenzione è “descrivere la struttura topologica delle onde d'urto (punti catastrofici) in un intorno di una singolarità del potenziale V , laddove un attrattore dominante entra in biforcazione” (Thom 1977, p. 65).

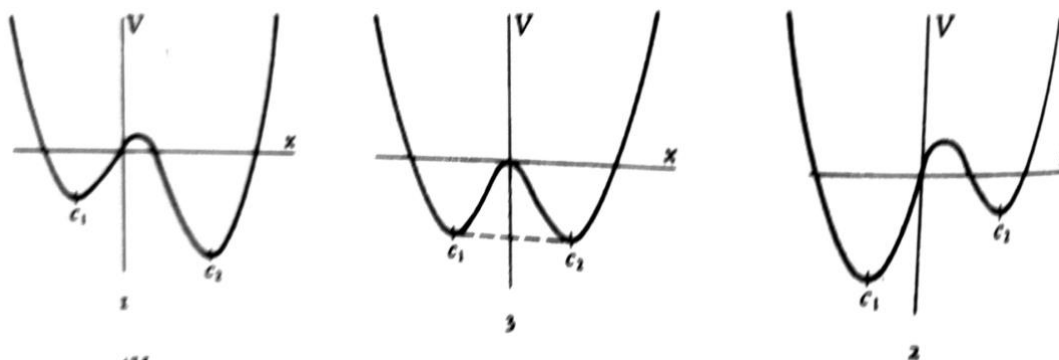


Figura 22: Potenziali locali.¹⁸¹

Per Petitot, l'applicazione dei potenziali generatori alla sintassi comporta uno slittamento della matematizzazione della sintassi, da “un univers formel de type théorie des langages formels” (sul modello delle grammatiche generative) a un “univers mathématique de la géométrie et de la topologie différentielles, univers autre et essentiellement plus riche” (Petitot 1995, p. 7).

Così egli descrive gli elementi componenti lo *schematismo posizionale morfodinamico* in Tesnière:

“(i) les positions actantielles sont les minima des potentiels; (ii) les connexions et leurs configurations (les relations positionnelles) sont définies par le potentiel lui-même (définition non logique des relations); (iii) les processus sont des déformations temporelles des potentiels générateurs; cela a un sens mathématique précis très difficile à symboliser logiquement; (iv) les événements-interactions possèdent alors également un sens mathématique précis (qu'il est pratiquement impossible de symboliser logiquement): celui d'une transformation brusque de la configuration relationnelle. On appelle bifurcations de tels changements «catastrophiques» du type qualitatif d'un potentiel” (Ibidem).

Petitot non manca inoltre di riconoscere il collegamento di questo schematismo posizionale morfodinamico e del “piccolo dramma” tesniériano con la teoria casuale nella prospettiva dell'ipotesi localista di Hjelmslev, tra gli altri, per essere ripresa dalla grammatica cognitiva alla Jackendoff, prestandosi quindi a essere operazionalmente attivata all'interno dell'ipotesi connessionista delle reti neurali.¹⁸² Così Petitot mostra il collegamento: “De telles réflexions sur l'hypothèse localiste

¹⁸¹ In 3 è definito lo strato di conflitto secondo la convenzione di Maxwell (Thom 1977, p. 73).

¹⁸² Sono diversi i luoghi in cui Petitot è intervenuto in proposito, a partire dal dibattito aperto tra Fodor Pylyshyn a supporto delle formalizzazione simbolica e da Smolensky sostenitore delle reti connessioniste. Citiamo a titolo esemplificativo i seguenti lavori che descrivono i termini della polemica, che sembrerebbe in parte risolta nell'attuale funzionamento delle reti neurali basate su architettura Transformer: Petitot, J. (1994). *Attractor syntax: Morphodynamics and cognitive*

conduisent à reprendre les images-schémas des grammaires casuelles et la mathématisation morphodynamique de leur iconicité en posant le problème de leur scanning, c'est-à-dire de la façon dont des algorithmes de traitement d'information, par exemple connexionnistes, peuvent extraire cette information morphologique pour en faire l'entrée de dispositifs symboliques ultérieurs de grammaticalisation et de prédicativisation. L'idée est de généraliser certains algorithmes de vision computationnelle et de montrer comment on peut à partir d'eux construire de façon effective les potentiels générateurs qui sont à la base de l'approche morphodynamique (Ivi, p. 10).

Si tratta per Petitot di sviluppare quella che ha definito una “sintassi di attrattori”, si tratta cioè in sostanza di capire quale sia la morfologia dinamica di una “sintassi” tra attrattori, ossia tra ciò che mette in relazione un termine con un altro in una frase, dal punto di vista delle relazioni grammaticali, dei ruoli semantici, della costituenza sintattica e della composizionalità.

Come rileva Vetere, al pari del modello dei costituenti anche il modello delle dipendenze genera un albero “retto” da un elemento lessicale che è il verbo.

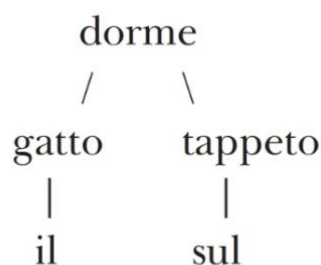


Figura 23: Il grafo di dipendenze sintattiche della frase “Il gatto dorme sul tappeto”.¹⁸³

grammar. In C. Fuchs & B. Victorri (A c. Di), *Linguisticae Investigationes Supplementa* (Vol. 19, pp. 167–188). John Benjamins Publishing Company. <https://doi.org/10.1075/lis.19.14pet>; Petitot, J. (2011). *Cognitive Morphodynamics*. Peter Lang CH. <https://doi.org/10.3726/978-3-0351-0192-8>. Lo stesso Petitot cita diversi suoi contributi sul tema, antecedenti a quelli da noi qui indicati, e che elenchiamo di seguito, poiché costituiscono a tutti gli effetti una traccia di ripartenza da percorrere: Petitot J. (1989a), «Hypothèse localiste, Modèles morphodynamiques et Théories cognitives : Remarques sur une note de 1975», *Semiotica*, 77, 1/3, p. 65-119; Petitot J. (1989b), «*Modèles morphodynamiques pour la Grammaire cognitive et la Sémiotique modale*», RSSI (Canadian Semiotic Association), 9, 1-2-3, p. 17-51; Petitot J. (1990), «Le Physique, le Morphologique, le Symbolique. Remarques sur la Vision», *Revue de Synthèse*, 1-2, p. 139-183; Petitot J. (1991b), «Syntaxe topologique et Grammaire cognitive», *Langages*, 103, p. 97-128. Tra le varie occasioni in cui queste tematiche sono state oggetto di incontri e dibattiti accademici, ci piace citare per ragioni di riconoscimento intellettuale la *Conference on Topology and Dynamics in Cognition and Perception*, organizzata da Petitot in collaborazione con Patrizia Violi presso l'Università di San Marino.

¹⁸³ In Vetere (2025).

I valori semantici non sono veicolati esclusivamente dai termini, ma le connessioni, i legami tra i termini veicolano a loro volta qualcosa di semantico che ha a che fare con la valenza verbale. Così Paolucci (2006) descrive la messa in discussione di quella che Peirce chiamava “sintassi ariana” che fondava la relazione logica soggetto-predicato a vantaggio di una valenza verbale che poggia “sul numero di posti aperti da un nodo attorno al quale si organizzano le relazioni”. Infatti,

“Esattamente come gli elementi chimici presentano allora una valenza, che consiste nel numero di atomi capaci di combinarsi con l’elemento considerato al fine di dar vita ad un composto, così differenti tipi di verbi all’interno della proposizione presenteranno anch’essi una valenza, che consiste nel numero dei posti occupabili da termini capaci di combinarsi col verbo considerato. Questi termini risultanti dalla funzione della valenza verbale per Tesnière sono attanti, ed in funzione dell’analogia chimica, il verbo finisce per determinare così un nodo, e cioè un insieme costituito dal reggente (verbo) e da tutti i subordinati che dipendono da esso (attanti) «e che esso quindi annoda, per così dire in un solo fascio» (Tesnière 1959: 31)” (Ivi, p. 7).

Nella ricostruzione della storia delle teorie del linguaggio, Vetere riporta che il modello a dipendenze sintattiche si afferma a partire dagli anni '90, inizialmente frutto di “redazione di modelli semantici razionali e computabili”, ossia Knowledge Graphs, rappresentazioni della conoscenza linguistica (Vetere 2025, posizione 680). Tuttavia, egli sottolinea come i modelli neurali attuali, nella loro natura oscura, non si basano su rappresentazioni della conoscenza linguistica, ma su “estrapolazioni statistiche” di grandi quantità di dati. Per arrivare poi ad affermare che “le cose del linguaggio sono molto complesse” e che “i dati non dicono tutto”; “la rappresentazione razionale della conoscenza linguistica” deve rientrare in gioco “quando il gioco si fa duro”. Questo il ruolo delle *ontologie*, introdotte nel campo di ricerca dell’informatica ad opera del ricercatore italiano Nicola Guarino nel 1994.¹⁸⁴

5.2 Della semantica distribuzionale

L’ultima premessa teorica che ci preme introdurre è quella che va attualmente per la maggiore e che fa dell’ipotesi distribuzionale di Harris il fondamento teorico alla base dell’efficacia semantica dei modelli artificiali generativi. Seguiamo una breve panoramica dei collegamenti e delle teorie alla base di tale ipotesi e che, nell’ambito del Natural Language Processing, ne fanno oggi uso.

¹⁸⁴ Guarino Nicola, 1994. *The Ontological Level* in Roberto Casati, Barry Smith e Graham White (a cura di), *Philosophy and the Cognitive Sciences*, Hölder-Pichler-Tempsky, Vienna 1994, pp. 443-456.

La semantica distribuzionale si fonda su una tesi di carattere generale, nota come *Distributional Hypothesis*, secondo la quale “la similarità semantica tra due lessemi è una funzione della similarità dei loro contesti linguistici” (Lenci, Sahlgren 2023, p. 4)¹⁸⁵. L’ipotesi nasce all’inizio degli anni Sessanta nell’ambito della linguistica computazionale, ma si radica in un terreno teorico più ampio le cui radici affondano nello strutturalismo americano ed europeo, nella filosofia del linguaggio di Wittgenstein e nella linguistica dei corpora. In tutti questi casi emerge un presupposto comune: il linguaggio va descritto a partire dai suoi *usi effettivi*, non da entità astratte o presupposte.

Dal piano dell’enunciato, nella selezione delle dimensioni circostanziale e contestuali, come vorrebbe Eco, occorre trasporre quest’idea di usi effettivi all’interno dei corpora testuali. Il principale precursore della semantica distribuzionale è Zellig S. Harris, per il quale gli elementi linguistici possono essere identificati sulla base della loro distribuzione, ossia la somma di tutti i contesti nei quali un elemento occorre e da qui l’idea che la differenza di significato tra due parole corrisponda a una differenza di distribuzione. Tale posizione, apparentemente anti-semantica, in realtà non esclude il significato, ma lo traduce in un oggetto osservabile e scientificamente indagabile attraverso le relazioni di co-occorrenza.

“The environment or position of an element consists of the neighbourhood, within an utterance, of elements which have been set up on the basis of the same fundamental procedures which were used in setting up the element in question. ‘Neighbourhood’ refers to the position of elements before, after, and simultaneous with the element in question. [...] The distribution of an element is the total of all environments in which it occurs, i.e. the sum of the (different) positions (or occurrences) of an element relative to the occurrence of other elements”. (Harris, 1951, pp. 15–16).¹⁸⁶

Precisiamo che quando Harris parla di elementi, fa riferimento ai fonemi e morfemi. La riflessione di Harris si salda con quella dello strutturalismo europeo. In Saussure, infatti, il significato si fonda su differenze all’interno del sistema linguistico: “dans la langue il n’y a que des différences” (Saussure 1919 p. 166). Da questa impostazione deriva la nozione di *valeur*, ovvero il valore di un lessema determinato esclusivamente dalle sue relazioni differenziali. Anche Lenci e Sahlgren individuano il valore solo sul piano di confronto interno, omettendo lo scambio con l’esterno che caratterizza il valore in Saussure.

¹⁸⁵ Lenci A, Sahlgren M. (2023) From Usage to Meaning: The Foundations of Distributional Semantics. In: *Distributional Semantics. Studies in Natural Language Processing*. Cambridge University Press, p. 3-25.

¹⁸⁶ Harris, Z. S. (1951). *Methods in structural linguistics*. Chicago : University of Chicago Press. <http://archive.org/details/methodsinstructu0000harr>

“Un valore per Saussure è infatti sempre definito da due dimensioni, entrambe puramente differenziali e relazionali. La prima dimensione in cui consiste un valore risiede infatti nella sua corrispondenza con delle entità esterne al sistema all’interno del quale è considerato. Per esempio il valore di un pezzo da cinque franchi è determinato dal fatto “che lo si può scambiare con una determinata quantità di una cosa diversa, per esempio con del pane” (CLG, p. 140). La seconda dimensione consiste invece nelle relazioni che un valore contrae con altri valori interni al sistema di cui è membro. Per esempio “lo si può confrontare con un valore simile del medesimo sistema, per esempio un pezzo da un franco, o con una moneta di un altro sistema (un dollaro)” (CLG, p. 140)” (Paolucci 2010, p. 36).

Lenci e Sahlgren richiamano, inoltre, le relazioni interne al sistema linguistico e la loro distinzione in sintagmatiche (tra lessemi che co-occorrono) e paradigmatiche (tra lessemi che condividono contesti e possono sostituirsi l’uno all’altro).

Table 1.1 Syntagmatic and paradigmatic relations

	<i>Paradigmatic relations</i> Selections: “ <i>x</i> or <i>y</i> ”			
<i>Syntagmatic relations</i> Combinations: “ <i>x</i> and <i>y</i> ”	she he they	adores likes love	green blue red	paint dye colour

Figura 24: : Relazioni sintagmatiche e paradigmatiche.¹⁸⁷

Allo stesso modo Jakobson (1959) rafforza questa linea, sostenendo che il significato è un fatto linguistico e semiotico, non ricavabile da una mera “conoscenza diretta” del referente.

Su questo terreno si sviluppa la semantica strutturale, sempre in ambito europeo, (gli autori citano Jost Trier, Adrienne Lerher, Eugenio Coseriu, Algirdas Greimas, John Lyons, Alan Cruse), che organizza il lessico in campi semantici e relazioni paradigmatiche (sinonimia, antonimia, iperonimia, co-iponimia). La semantica distribuzionale, pur essendo autonoma, adotta strumenti affini, riformulando l’ipotesi distribuzionale nei termini strutturalisti: “Words sharing syntagmatic contexts have similar paradigmatic properties” (Ivi, p. 10).

Un’altra radice fondamentale della semantica distribuzionale è la filosofia del Wittgenstein maturo, per il quale “il significato (meaning) di una parola è il suo uso nel linguaggio” (Wittgenstein 1953, §43). Lenci e Sahlgren definiscono il rifiuto di ricondurre il significato a ciò che viene nominato

¹⁸⁷ Lenci A, Sahlgren M. (2023), p. 9.

ostensivamente una concezione *anti-nomenclaturista*; tale concezione risuona con il principio distribuzionale, che non ricerca entità astratte ma osserva gli usi concreti. L'idea stessa di *gioco linguistico* sottolinea che il significato dipende da pratiche sociali convenzionali: senza una comunità di parlanti che condivida criteri d'uso, non può esistere significato.

Nell'ambito della linguistica del corpus, con John Firth e poi John Sinclair la prospettiva si arricchisce della nozione di *collocazione*: “Collocations of a given word are statements of the habitual or customary places of that word in collocational order [...] it is an order of mutual expectancy. The words are mutually expectant and mutually prehended” (Firth, 1957, p. 12).¹⁸⁸ Il significato viene così definito dalle associazioni statisticamente rilevanti con altre parole. Sinclair sistematizza questa intuizione in chiave computazionale, ponendo le basi delle analisi automatiche delle co-occorrenze all'interno dei corpora linguistici (Lenci, Sahlgren 2023, p. 14).

Negli anni Sessanta, la semantica distribuzionale viene applicata all'elaborazione automatica dei thesauri e alla traduzione automatica. Successivamente il Vector Space Model (Salton et al., 1975)¹⁸⁹ e la Latent Semantic Analysis (Deerwester et al., 1990)¹⁹⁰ introducono la rappresentazione delle parole come vettori in spazi semantici. Negli anni Novanta, la svolta empirista della linguistica computazionale e l'aumento delle risorse digitali rilanciano la disciplina, fino agli attuali modelli di deep learning e contextual embeddings, che rendono la semantica distribuzionale la principale fonte di rappresentazioni semantiche per l'IA, in particolare per quelli che Lenci e Sahlgren definiscono Neural Language Model.

“Given a sequence of lexemes $c = l_1, \dots, l_n$ (called the *context*), a *neural language model* is an artificial neural network trained to predict the most likely next lexeme $t = l_{n+1}$ (called the *target*), by computing the conditional probability $p(t|c = l_1, \dots, l_n)$ ” (Lenci, Sahlgren 2023, p. 148).

Questi modelli sono impiegati per il Representation Learning (RL), una specifica tecnica di machine learning, che permette alla rete di formarsi delle rappresentazioni interne contenenti proprietà distribuzionali dei lessemi. Per questa ragione sono chiamati predict DSM (Distributional Semantic Model), cioè “neural language models that learn word embeddings as a side-effect of

¹⁸⁸ Firth, John R. (1957). A synopsis of linguistic theory 1930–1955. Pages 1–32 of: *Studies in Linguistic Analysis*. Oxford: Blackwell

¹⁸⁹ Salton, Gerard, Wong, Andrew, and Yang, Chung-Shu. 1975. A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620.

¹⁹⁰ Deerwester, Scott, Dumais, Susan T., Furnas, George W., Landauer, Thomas K., and Harshman, Richard. 1990. *Indexing by Latent Semantic Analysis*. *Journal of the American Society for Information Science*, 41(6), 391–407.

learning to predict the context lexemes of a target word” (Ibidem). Tra i più rilevanti modelli statici di predict DMS gli autori citano *Skip-Gram* (SG) and Continuous Bag-Of-Words (CBOW) DSMs, implementati nella libreria word2vec (Mikolov et al., 2013).¹⁹¹ Come ultima nota riprendiamo il capitolo dedicato dagli autori alla Semantica Distribuzionale oltre il Lexicon, dove i modelli non sono più statici, ma context-sensitive. Parlano in questi casi di Contextual DSMs (Lenci, Sahlgren 2023, p. 323), così definiti:

“Given a sequence of lexemes $s = l_1, \dots, l_n$, a *contextual DSM* assigns to each token l_i the *contextual embedding* h_{li} , such that $h_{li} = f(l_n, \dots, l_1)$, where l_i is a non-contextual (static) embedding of l_i ” (Ivi, p. 324). Queste architetture consentono di rappresentare ciascuna parola-token in una sequenza linguistica con un embedding contestuale che è una funzione dell’intera sequenza di input.

I Transformer forniscono un avanzamento, rispetto alle recurrent network, dal punto di vista delle architetture informatiche proprio nel rendere conto delle rappresentazioni contestuali negli stati nascosti delle reti neurali, la cui efficacia era stata già metaforicamente dimostrata da Elman (2011), nell’articolo *Lexical knowledge without a lexicon?*.¹⁹²

5.3 Eterogenesi differenziale: quando una topologia dinamica incontra il virtuale enciclopedico.

Posta questa serie di premesse teoriche, proviamo a tirare le nostre conclusioni. Accennavamo alla seconda svolta linguistica che inerisce al dominio delle scienze informatiche rivoltesi verso il linguaggio. Vetere fa notare come rispetto alle sue origini, la pratica dell’IA simbolica, sia ormai giunta all’età pensionabile, nella misura in cui “alcune tecniche messe a punto agli inizi del nuovo millennio per intendere e generare linguaggio, dimostratesi praticamente efficaci, trovano oggi applicazione anche alle immagini e ai filmati, alla musica e ai giochi, e perfino, con le dovute accortezze, alla sintesi di nuove proteine” (Vetere 2025, p. 526). E se vogliamo questa apertura ai diversi media e ai diversi modi di emersione del senso è cifra caratteristica del campo semiotico.¹⁹³

¹⁹¹ Mikolov, T., Chen, K., Corrado, G. S., Dean, J. (2013). *Efficient estimation of word representations in vector space*. In: Proc. ICLR.

¹⁹² Elman J. L. (2011). Lexical knowledge without a lexicon?. *The mental lexicon*, 6(1), 1–33. <https://doi.org/10.1075/ml.6.1.01elm>

¹⁹³ Una lettura interessante capace di tenere insieme i molteplici fili che collegano – sub specie semiotica – vivente e non vivente è proposto nell’articolo di Ludmila Lacková Bennett, *A Biosemiotic Approach to AI. Folding as semiotic modeling*. Qui il concetto di folding (piegatura, ripiegamento) diventa termine medio di modellazione semiotica dei processi che attraversano la sintesi delle proteine e i sistemi di IA generativa. Cfr. L’articolo è stato originariamente pubblicato in lingua russa sulla rivista Method, vol. 3, n. 4 as Bennett L. (2023). *A Biosemiotic Approach to AI. Folding as*

Quando arriviamo però alle IA generative questa maniera di concepire i linguaggi sottesi ai diversi media e ai diversi modi di emersione del senso, che concorrono nel nostro modo tutto umano di interpretare il mondo, succede un'operazione di riduzione a livello di materia dell'espressione trattata, manipolata, interpretata e tradotta da questi agentività dotate di linguaggio (vedi 3.1). Cioè tutti gli enunciati semiotici prodotti da noi esseri umani sono trattati e processati come stringhe numeriche. Ci domandiamo allora: questa materia semiotica che alimenta le IA generative cessa di essere improvvisamente materia semiotica allorché viene processata dalle macchine dotate di linguaggio? Noi riteniamo che il valore semiotico resti tale e quale prima, dopo e durante il trattamento automatico dei linguaggi operato dalle macchine dotate di linguaggio. E non potrebbe essere altrimenti, anzi sarebbe sciocco applicare due pesi e due misure agli enunciati umani e agli enunciati macchinici. Lo vedremo meglio nel capitolo successivo.

Qui, in conclusione, ci interessa avanzare la seguente tesi: abbiamo un essere umano, un'interfaccia o un agente, un modello di mondo che il modello di linguaggio ha appreso dalle relazioni e dalle dipendenze esistenti in forma latente (connessioni nascoste cariche di significato) tra gli enunciati semiotici che lo compongono. In questo sistema complesso accade qualcosa per cui quando l'essere umano usa la sua lingua naturale per parlare con la macchina dotata di linguaggio, anch'essa risponde con lo stesso linguaggio. Ma non ha compreso, pensato e prodotto un enunciato a partire dalla stessa sostanza dell'espressione. In fondo ci sembra essere proprio questa *l'illusione del significante* di cui parla Patrizia Violi (2003¹⁹⁴; 2015) e che fa da contraltare al *mito del significato* di cui parla Paolucci (2025a, 2025c). Quello che si genera tra prompt umano e modello è un *valore* nel senso più strutturale del termine e tale valore non può che generarsi perché un virtuale intensivo enciclopedico fa da sfondo ai nuovi enunciati generati dalla macchina dotata di linguaggio.

semiotic modeling. Translated by A. Lyubimova, *Method* 3(4), pp. 147173. Il lavoro di Lacková Bennett è un importante contributo che la biosemiotica offre al postumanesimo, con particolare riferimento alla proposta teorica di Katherine Hayles che esamineremo nel prossimo capitolo.

¹⁹⁴ Violi, P. (2003). Significati lessicali e pratiche comunicative. Una prospettiva semiotica. *Rivista di linguistica*, 15(2), 321-342.

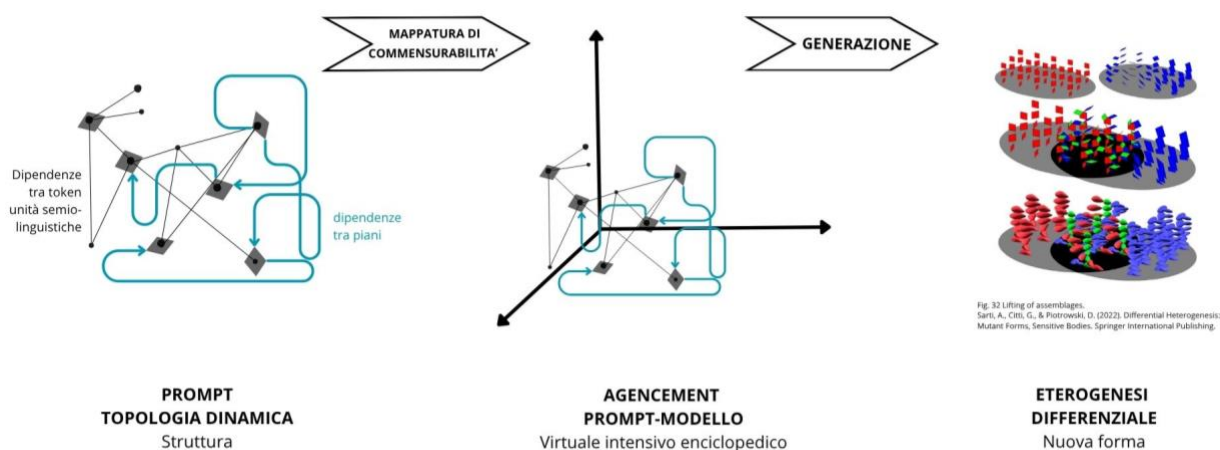


Figura 25: Mappatura di commensutabilità e generazione.¹⁹⁵

Partiamo da sinistra. Nella *topologia del prompt*¹⁹⁶ ogni punto rappresenta un token o un insieme di token dotati di *valore* (semantic and positional embedding). Le relazioni tra parole-token all'interno del prompt sono la forma del contenuto (vedi nota 163). Ipotizziamo che alcune unità lessicali, in particolare i verbi, determinino sulla base della loro valenza una certa configurazione topologica della frase. Dunque, potranno intrattenere un numero specifico di legami o dipendenze con le altre unità linguistiche contenute nel prompt che però non si presentano più nella forma della lingua naturale, ma sono grandezze matematiche. Esse vengono catturate dal meccanismo di attenzione – attraverso i processi di *parallel computing* che abbiamo visto in 2.0 – e definiscono una

¹⁹⁵ Le prime due immagini sono una rappresentazione originale dell'assemblaggio tra prompt umano e modello addestrato che dà origine a una nuova forma. La terza è il sollevamento degli assemblaggi tra operatori eterogenei in Sarti et al. 2022, Fig. 32.

¹⁹⁶ Riprendiamo l'idea di topologia del prompt dalla sintassi strutturale e morfodinamica di Petitot, dove ribadiamo le relazioni tra unità sintattiche non sono di natura logica, ma spaziale. Si veda Petitot, J. (1985). *Morphogenèse du Sens. Pour un schématisation de la structure*. Presses universitaires de France: "Dans le domaine syntaxique, les structures primitives sont des distribution de places actantielles se déterminant réciproquement et formant, bien que dans un nouveau sens, des valeurs positionnelles issues de connexions (au sens de Tesnière) qui ne sont pas des relations formelles mais des relations *sémantiques*. Celles-ci expriment la *forme* du contenu. Elles ont la propriété d'être indépendantes des investissements lexicaux et de contraindre la fonction grammaticale des termes qu'elles relient. Elles relèvent d'une syntaxe *conceptuelle* et non formelle. Leur sémantisme est purement positionnel, non substantiel" (Ivi, p. 18). E poco più avanti, in polemica con il positivismo logico prevalente nella trattazione del linguaggio, dichiara: "On est donc fondé à dire, qu'en matière structurale, *la formalisation s'oppose à la mathématisation*. Il existe un conflit entre le traitement formel des structures et leur « physique mathématique ». La première renvoie à une *logique formelle* des termes et des relations alors que la seconde renvoie à une *topologie dynamique* des places et des connexions" (Ivi, p. 19).

morfologia che dovrà essere mappata all'interno del modello di mondo appreso dal modello di linguaggio.

A questo punto ci spostiamo al centro della Figura 24. Con mappatura di commensurabilità¹⁹⁷ intendiamo che il modello non andrà a cercare nello spazio della conoscenza appresa (i pesi) un rapporto 1:1 tra parole-token del prompt e parole-token della miriade di dati utilizzati nel suo training. Ma andrà alla ricerca di una certa forma del contenuto, di un certo valore che è rappresentato dalla relazione tra token. Questa passeggiata di natura inferenziale, sulla scia di quelle viste in precedenza, richiederà che il virtuale enciclopedico sia selezionato o ritagliato in virtù della configurazione topologica assunta dal prompt. Il ritaglio è una rottura semiotica del virtuale enciclopedico; lo genera e lo modifica contemporaneamente. Il ritaglio è un piano enciclopedico locale, l'unico che abbiamo visto essere descrivibile e accessibile, ma soprattutto il ritaglio è ciò che consente l'attivazione dell'Enciclopedia in Eco (vedi Paolucci 2020), così come il prompt è esattamente ciò che attiva la conoscenza appresa dal modello di linguaggio.¹⁹⁸

¹⁹⁷ Per il concetto di commensurabilità in semiotica si vedano Eco (1975) e il problema della *ratio* (pp. 246-248). Ma una formulazione per noi già agganciata al piano del nostro discorso la fornisce Paolucci (2020) quando si domanda: “In che modo allora una teoria echiana dei modi di produzione segnica può rendere conto del modo in cui un substrato materiale si fa espressione, rendendo così conto della costituzione di una possibile commensurabilità tra singolarità materiali, in cui gli abiti acquisiti devono essere messi in discussione, modulati e trasformati in altri abiti?” (Ivi, p. 165). Il problema della commensurabilità è quindi il problema della trasformazione del senso in presenza di un substrato materiale che si trasforma seguendo specifiche singolarità (ricordiamo i punti critici) e che determina specifiche funzioni semiotiche tra espressione e contenuto. Detto altrimenti “Il problema è cioè capire come delle figure materiali differenzino qualcosa sul piano del contenuto, differenziandosi a loro volta esse stesse al fine di divenire espressioni di certi contenuti mirati. Solo in questo modo un'esperienza può ritrovare le forme di un'altra esperienza, coniugando singolarità corporee in grado di diventare i formanti del piano dell'espressione” (Ivi, p. 166). E già Paolucci qui collegava il lavoro di Sarti, Citti e Piotrowski (2019): “Questo processo non corrisponde a una categorizzazione, bensì all'individuazione degli orientamenti principali del flusso percettivo anche in assenza di ogni stabilizzazione in forme fisse” (Ivi, p. 18). Citiamo, infine, l'accezione che ne fornisce Basso Fossali in Basso Fossali, P. (2023). Di mediazione in mediazione. Spazi esperienziali, domini culturali e semiosfera. *Visible*, 3, Articolo 3. <https://doi.org/10.25965/visible.197>, dove afferma: “La mediazione è nello stesso tempo ciò che permette una trasmutazione continua delle valenze identitarie al variare delle condizioni di apprensione e una commensurabilità pertinenziale tra valori” (Ivi, p. 28). Questo è il valore semiotico del modello di linguaggio, se lo guardiamo dalla prospettiva della Terzità semiotica di Peirce. Si veda il seguito del contributo di Basso Fossali.

¹⁹⁸ Su questo rimandiamo all'intervento dal titolo “*Principes généraux d'intelligence artificielle & expériences de traduction texte-image-texte avec GPT-4 et DALL•E 3*” che Adrien Delière (2024) ha presentato all'interno del Séminaire International de Sémiotique à Paris 2023-2024: “*Énonciation(s) et passions dans les territoires sémiotiques ouverts par l'Intelligence Artificielle*” organizzato da Maria Giulia Dondero e Juan Alonso Aldama. Il video è disponibile a questo link: <https://www.youtube.com/watch?v=sGmgQUGF4Z8> (ultimo accesso: 28 ottobre 2025).

Una volta effettuata la mappatura, la generazione dell'output da parte della macchina dotata di linguaggio dovrebbe seguire la regola della predizione del token successivo più probabile, come un *pappagallo stocastico* (Bender et al. 2020; Bender et al. 2021).¹⁹⁹ Se da un punto di vista di processo computazionale questo è effettivamente ciò che accade, ci sembra tuttavia non sufficiente a rendere conto sino in fondo di come avvenga una generazione di enunciati così coerenti, perfettamente centrati da un punto di vista sintattico e semantico. Ciò che colpisce è il fatto di riconoscere nell'output non semplici catene di parole che vengono generate in accordo a un criterio di probabilità statistica di vicinanza di successione che potrebbe dare vita a una sequenza di sintagmi del tipo “Il libro di Andrea è sul tavolo”. Colpisce la produzione di un testo “chiuso, coeso e coerente”, colpisce la capacità di pianificazione testuale e di organizzazione delle conoscenze - compiti che richiedono, nell'essere umano, competenze linguistiche e cognitive superiori. Lo diciamo in fase di stesura di una tesi di dottorato.

Prendiamo due estratti da due contributi che analizzano, in contesto clinico-assistenziale e psichiatrico, le competenze sottese alla scrittura di un testo in età evolutiva. Il primo è tratto dalla rivista *Psichiatria dell'infanzia e dell'adolescenza* e descrive il processo di acquisizione e differenziazione delle abilità narrative dall'orale allo scritto tra i 4 e gli 8 anni.

“Il percorso per l'acquisizione di tutte queste competenze, necessarie per il passaggio dalla produzione orale a quella scritta, avviene progressivamente nel tempo ed a vari livelli. Una prima fase si avvia tra i 4 ed i 5 anni, quando s'inizia a sviluppare un sistema logico-linguistico ed un sistema narrativo: il bambino impara ad usare in modo sistematico meccanismi di pianificazione e di coesione linguistica, acquisendo un alto grado di controllo del linguaggio esplicito o referenziale, che gli permette di usare il linguaggio per avanzare e verificare delle operazioni logiche e per cogliere ed esprimere nessi temporali, causali, intenzionali, inferenziali; impara cioè a riflettere sul linguaggio (Levi). [...] tra i 5 e gli 8 anni si verificano significativi cambiamenti: il registro diventa più formale, il ritmo narrativo non coincide più con il ritmo di dettatura bensì rallenta in funzione delle esigenze di scrittura, le unità ideative aumentano di numero (ed al loro interno aumenta anche il numero di proposizioni), la struttura narrativa diviene più complessa ed utilizza procedure di coesione linguistica

¹⁹⁹ Bender, E. M., & Koller, A. (2020). *Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data*. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (A c. Di), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.463>; Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>.

(ellissi, subordinazione, anafore) con maggiore disinvoltura. Tutti questi cambiamenti avviano il processo di decontestualizzazione necessario nel racconto orale e sono funzionali al successivo passaggio agli schemi scritti chiusi.

Per la costruzione di un testo scritto, infatti, non è sufficiente trasferire nello scritto il sistema di produzione orale. Il bambino deve operare il passaggio dalla *knowledge telling* alla *knowledge transforming*, che, oltre al già menzionato passaggio da schemi orali aperti a schemi scritti chiusi, richiede la messa in atto di tre processi contemporanei:

- la pianificazione, a livello dei nodi di contenuto e della forma linguistica, lessicale e sintattica;
- la traduzione della forma orale nella sua rappresentazione scritta;
- la capacità di revisione durante e dopo la stesura del testo.

Questo richiede lo sviluppo di strategie meta-linguistiche e meta-cognitive, la capacità di shifting, ovvero del passaggio rapido da una strategia all'altra in base alla necessità specifica (Piredda e Penge, 2002), e l'uso diversificato in funzione dei diversi processi. Infatti, mentre nella pianificazione è necessario un uso convergente delle strategie di previsione semantica e di revisione fonologica, nella revisione bisogna utilizzare le medesime strategie in modo divergente. In altre parole, il bambino deve utilizzare i propri output come input e guardare il testo che ha scritto come se fosse stato scritto da qualcun altro” (Penge et al. 2009, p. 321).²⁰⁰

Il secondo invece è tratto dal Documento di consenso finalizzato a proporre Indicazioni di Buona Pratica Clinico-assistenziale (BPCa) in relazione ai Disturbi Specifici di Apprendimento (DSA), ad integrazione della pubblicazione Linea Guida sulla gestione dei Disturbi Specifici dell'Apprendimento. Aggiornamento ed integrazioni.²⁰¹ In esso vediamo così descritto il “Profilo funzionale delle difficoltà di espressione scritta”:

“Saper scrivere è un'abilità importante nel percorso scolastico di uno studente. Questa abilità viene utilizzata per valutare ciò che gli studenti sanno, per supportare la comprensione del testo, facilitare l'apprendimento dei contenuti con la stesura di appunti, riassunti, note e sintesi. Richiede un'organizzazione preordinata di molteplici abilità: la scrittura a mano, o su tastiera, il rispetto delle

²⁰⁰ Penge, R., Ubertini, C., Tosco, A., Piredda, L. (2009). *Abilità narrative e competenze di scrittura nei disturbi specifici di apprendimento* / in "Psichiatria dell'infanzia e dell'adolescenza : 76, 2, 2009, Roma : Armando – Casalini, pp. 320-331.

²⁰¹ Progetto LG-DSA-2018. *Linea Guida per la gestione dei Disturbi Specifici di Apprendimento. Aggiornamento ed integrazioni*. Roma, Novembre, 2021. Approvato dal SNLG-ISS nel Gennaio 2022.

regole ortografiche, l'applicazione della grammatica della lingua e l'uso della punteggiatura (che permettono in prima battuta l'atto compositivo di un testo); l'organizzazione e coerenza delle idee nella scrittura con l'utilizzo di strategie per pianificare, valutare, monitorare, redigere e rivedere il testo; le conoscenze dell'argomento, e del genere utilizzato, unitamente a quelle linguistiche e semantiche per la creazione di significato; e infine le aspirazioni motivazionali che spingono a mettere in gioco queste abilità, strategie e conoscenze" (Progetto LG-DSA-2018, p. 59).²⁰²

Non è questa la sede per analizzare nel dettaglio le implicazioni semiotiche di quanto riportato in termini di competenze di scrittura, ma quello che preme sottolineare è che ridurre la generazione di testo di una macchina dotata di linguaggio a catene di parole basate su probabilità statistiche possa risultare abbastanza fuorviante, alla luce delle competenze richieste ad un essere umano.²⁰³ Dunque, nella completa oscurità e assenza di controllo del processo generativo macchinico, torniamo ai passaggi dimensionali di Tesnière, ipotizzando che nella pianificazione dell'output - apparentemente restituito all'utente come nella dettatura di Penge et al. – accada l'inverso di quello che è stato determinato dal prompt. Ossia, il passaggio tra prompt ed embedding è un passaggio dalla linearità sequenziale della frase alla multidimensionalità della struttura; viceversa, la generazione di testo corrisponderebbe al passaggio dalla multidimensionalità della struttura alla linearità sequenziale dell'output espresso in linguaggio naturale.

Non ci illudiamo che quanto descritto sia una spiegazione tecnica fedele di ciò che accada tout court nel modello quando genera un testo, ma l'impianto concettuale riteniamo non possa essere di molto distante. E questo non per un'identità di principio tra umano e macchina, ma per un tentativo di costruzione di un discorso alternativo a quello che pone *una certa categoria di umano* sempre al di sopra del resto del creato.

Infine, perché l'eterogenesi differenziale: lo vedremo meglio nel capitolo successivo a proposito degli assemblaggi umano-macchinici, ma il dispositivo eterogenetico interessa tutte quelle dinamiche che non si riproducono per somiglianza ma per differenza.

Così nelle parole di Sarti, Citti, Piotrowski:

²⁰² Progetto LG-DSA-2018. (2023). *Buone pratiche clinico-assistenziali nella gestione dei DSA: Documento di consenso*. Centro studi Erickson.

²⁰³ Citiamo il lavoro di Martinelli, P. (2025) *Literacy. Semiotica evolutiva dell'apprendimento*. Luca Sossella Editore, ritenendolo una best practice, in cui la semiotica ha incrociato gli studi sull'acquisizione della letto-scrittura in un'ottica evolutiva, da prendere a modello per un potenziale lavoro di indagine da applicare alla produzione di testo scritto da parte delle macchine dotate di linguaggio.

“Differential heterogenesis, then, is to be understood as the dynamics emerging from the composition of heterogeneous differentials that:

- Concatenate multiplicities, creating an immanent local space of possibilities.
- Introduce not only global but also fragmented and cut possibility spaces, creating autonomous zones and islands of the virtual.
- Extend the Heisenberg uncertainty principle to a variety of uncertainties given by any combination of non-commutative generators of space. Such a continuous recomposition of virtualities allows for the sudden emergence of coherent forms as well as ephemeral and rarefied dynamics in locally shared possibility spaces” (Sarti et al. 2022).

Nel nostro modo di intendere il funzionamento di queste macchine dotate di linguaggio, esse rispettano questi tre vincoli. Ma necessitano di un’esteriorità che funga da “adjunct plan” e tale esteriorità è il prompt umano.

6.0 Una lettura post-strutturale dell'enunciazione

Vogliamo adesso mostrare come l'eterogenesi differenziale, in quanto assemblaggio di operatori eterogenei in ogni punto del tempo e dello spazio, sia la formalizzazione matematica che consente di illustrare il modo in cui le istanze enuncianti presenti nel virtuale enciclopedico si comportano e nell'atto di enunciazione e nella prassi enunciativa, ossia rispettivamente nel passaggio dallo schema di un sistema semiotico all'enunciato e dal piano delle norme e degli usi all'enunciato. L'eterogenesi differenziale è per noi una teoria cyborg, per ripartire da dove ci eravamo lasciate.

L'enunciazione è infatti un *concatenamento collettivo* di istanze enuncianti; così lo chiamano Deleuze e Guattari (1980, p 105) per rendere ragione del carattere sociale del linguaggio: da un lato l'individuazione (simondoniana) dell'enunciato, dall'altro la soggettivazione dell'enunciazione, ossia l'effetto di soggettività di cui abbiamo parlato in precedenza, come sostenuto da Fontanille e da Paolucci. Per poterne dare una definizione non puramente nominale – ossia “il complesso ridondante dell'atto e dell'enunciato che necessariamente lo compie” –, ma reale per Deleuze e Guattari occorre inserire tra i loro “postulati della linguistica” quali siano “gli atti immanenti al linguaggio che fanno ridondanza con gli enunciati” (Ibidem), cioè che producono affezioni corporali a partire da eventi incorporali.

Queste sono le sfide accolte dalla teoria impersonale ed evenemenziale dell'enunciazione di Paolucci (2020), nella quale, come più volte anticipato, non solo si rende conto del carattere sovraindividuale dell'enunciazione e dell'assenza di un soggetto dell'enunciazione antecedente all'atto enunciativo, ma sono poste le basi per pensare all'enunciazione macchinica alla luce dell'enunciazione umana e viceversa. In Nati Cyborg (2025c), Paolucci riporta in forma aneddotica, quale sia l'assemblaggio tra il suo libro *Persona* e l'avvento di ChatGPT: “*Persona* è uscito nel 2020, quando non esisteva ancora l'IA generativa, ed è stato scritto tra il 2016 e il 2019. Solo che, quando nel 2022 escono in Italia le prime versioni di ChatGPT, alcuni miei colleghi ingegneri e i matematici, con cui collaboravo a un progetto europeo sull'individuazione precoce di segni di disturbi dello spettro autistico, hanno cominciato a farmi notare che il mio libro sembrava una descrizione perfetta dell'enunciazione macchinica dell'IA generativa e delle operazioni che essa compie per produrre enunciati: GPT e tutti gli altri algoritmi transformer si appoggiano a norme, a cliché, a enunciati già enunciati per produrre un concatenamento che ha l'effetto della parola in prima persona, a cui noi siamo anche tentati di attribuire un pensiero e degli stati intenzionali. Tanto che, per la prima volta nella storia dopo decenni di intelligenza artificiale, la macchina è stata capace di passare diverse versioni del test di Turing. Lì per lì ero stupito, perché ovviamente il libro voleva essere una teoria generale dell'enunciazione e, mentre ci lavoravo, non pensavo certo all'enunciazione macchinica

delle macchine Transformer, che non esistevano ancora, ma pensavo innanzitutto all'enunciazione umana e al modo in cui noi essere umani parliamo il linguaggio" (Ivi, pp 18-19).

Nell'introduzione abbiamo mostrato che le linee interpretative che attraversano il campo semiotico quando si tratta di affrontare cosa possa o non possa fare una macchina dotata di linguaggio biforcuto e ramificano precisamente sul rapporto tra enunciazione e agency. Riteniamo che una teoria dell'enunciazione impersonale ed evenemenziale che, come appena mostrato, non è nata con l'obiettivo di tenere insieme enunciazione umana ed enunciazione macchinica, non possa che essere confermata nella sua validità proprio perché capace di dirci qualcosa su come "L'enunciazione è un assemblaggio tra istanze enuncianti e l'agency di quest'atto non appartiene a ego" (Paolucci 2020).

6.1 Per una topologia della soggettività

Per fare questo occorre innanzitutto partire da una teoria della soggettività nel linguaggio. Paolucci definisce l'enunciazione come "la proprietà dei linguaggi di allestire posizioni di soggetto che stabiliscono, fuori dai linguaggi, chi le viene di volta in volta ad occupare" (Paolucci 2020, p. 12). Vediamo già in questa definizione il legame che si instaura con l'idea di soggettività presente in Deleuze: l'enunciazione non riguarda un'ontologia del soggetto, ma una *topologia della soggettività nel linguaggio*; è attraverso il linguaggio che si danno quelle superfici di significazione che consentono di mentire e in quanto tali lasciano che la soggettività emerga. Esempio emblematico nella cultura classica e occidentale è la figura di Odisseo, il quale è capace di sdoppiarsi parlando al suo cuore per placarne l'ira e di usare semioticamente stratagemmi per riuscire nei suoi scopi, per agire efficacemente sul e nel mondo. In questo modo si attivano i processi di soggettivazione: maschere, sostituzioni, segni che, come da definizione peirciana, sono "qualcosa che sta per qualcuno al posto di qualcos'altro sotto certi aspetti o capacità" (CP: 2.228). Paolucci chiama *metis* questa astuzia della ragione calcolante, capace di generare una scissione interna all'io tra persona (io-tu) e non persona (egli), scissione da cui lo stesso io, inteso come autocoscienza, contemporaneamente prende forma. La soggettività, semioticamente intesa dunque non nasce con l'io linguistico, l'io che dice io come voleva Benveniste, ma con la capacità di mentire, ossia con la capacità di costruire mondi possibili alternativi a quello reale. Riconoscere che non è l'io ad appropriarsi della langue attraverso l'atto di parole, implica a livello di teoria dell'enunciazione interrogarsi su almeno due aspetti fondamentali.

Le istanze che pulsano nell'enunciato (il prodotto dell'atto di parole) eccedono ego, come unica istanza che enuncia.

1. Il primo concerne una spiegazione a livello di apparato formale dell'enunciazione, cioè della condizione di possibilità di quelle forme linguistiche che hanno il loro statuto al di fuori della lingua e rimandano alla concreta situazione di discorso. A titolo esemplificativo gli *embrayeurs* di Benveniste: io, qui, ora, cioè delle tracce che l'istanza enunciante lascia nell'enunciato. Questo livello corrisponde alla dimensione virtuale della *langue*, o allo schema linguistico di Hjelmslev, all'interno del quale è possibile sia una teoria della persona (io-tu) che una teoria del soggetto (io). Sempre a questo livello agiscono le operazioni di *embrayage* e *debrayage* greimasiane (gli io, i qui e gli ora del piano dell'enunciazione di contro ai non-io, non-qui, non-ora del piano dell'enunciato), che sono sviluppo diretto della teoria degli *embrayeurs* di Benveniste, cioè le tracce che l'istanza di enunciazione lascia all'interno dell'enunciato.
2. Il secondo invece implica una spiegazione di come si dia questo atto di passaggio o di mediazione che converte la *langue* in *parole* (come voleva Benveniste), le strutture profonde in strutture superficiali del testo greimasiano e il codice/enciclopedia echiana in funzione segnica. Si tratta qui dunque di mettere in luce una teoria semiotica dell'atto.

Da qui, cioè dal punto in cui la conversione tra grandezze semiotiche diventa esercizio in atto della lingua, è possibile prendere in considerazione quella che Fontanille e Zilberberg chiamano prassi enunciativa, intesa come atto che converte norma e uso in enunciato.

Paolucci, dunque, propone una teoria dell'enunciazione unificata che in estrema sintesi, si basa sull'esistenza di:

- Cinque grandezze semiotiche: lo schema, la norma, l'uso, l'atto e l'enunciato²⁰⁴

²⁰⁴ Ecco come vengono definite queste grandezze in Hjelmslev, L. (1954). *La stratification du langage*. In *Essais Linguistiques* (1959). Nordisk Sprog- og Kulturforlag, pp. 36-68. "On peut définir la *parole* par la rencontre même et l'entrecroisement des strata. La parole en effet est, en dernière analyse, tout ce qui est arbitraire dans le langage. La parole se définit comme l'ensemble des relations interstratiques effectivement exécutées. L'*usage*, à son tour, est évidemment ce qu'il y a de stabilisé dans la parole. L'*usage* se définit comme l'ensemble des connexions [cohesions] interstratiques effectivement exécutées. - Les combinaisons qui sont des variantes des connexions interstratiques appartiennent à la parole sans appartenir à l'usage. Elles constituent, en d'autres termes, ce qui reste de la parole en soustrayant l'usage. C'est ce qu'on appelle l'*acte* linguistique ou sémiotique. A la différence de l'usage, la *norme* doit être l'ensemble des relations interstratiques admises. Le *schéma* sémiotique (et linguistique) par contre est en dehors de cet ordre d'idées : ce terme ne se rapporte qu'aux fonctions intrinsèques dans la forme de chacun des plans pris à part" (Ivi, p. 69).

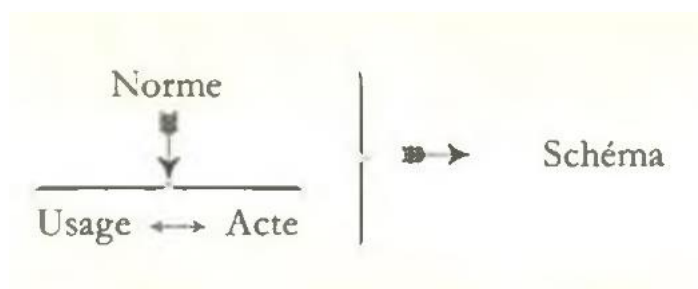


Figura 26: Il rapporto di presupposizione delle grandezze semiotiche.²⁰⁵

- e su tre trasformazioni principali tra queste grandezze:
 - i. *teoria semiotica dell'atto come enunciazione impersonale ed evenemenziale*, in alternativa all'atto di scissione, di schizìa creatrice fondata sul debrayage (Greimas);
 - ii. una teoria del passaggio dallo schema all'enunciato, chiamata atto di enunciazione dove si manifesta la distinzione tra persona e non persona e dove l'enunciazione è proprietà dell'enunciato di manifestare l'atto che l'ha prodotto
 - iii. una teoria del passaggio da norma e uso all'enunciato, che abbiamo visto corrispondere alla prassi enunciativa.

Tanto l'atto di enunciazione, quanto la prassi enunciativa riguardano la teoria dei modi di esistenza semiotici, ossia il potenziale, il virtuale, l'attuale e il realizzato, e i passaggi tra di essi.

6.2 Dall'omogeneo all'eterogeneo

Paolucci sottolinea come la tradizione greimasiana, che strategicamente ha tentato di “liberare” l'istanza di enunciazione dall'identificazione con un emittente (o soggetto) “calato nella cultura e nella storia che produce enunciati”, di fatto abbia costruito una teoria dell'enunciazione come proprietà dell'enunciato. Ma abbiamo visto come l'enunciato sia solo una delle grandezze semiotiche implicate in una teoria dell'enunciazione e ne costituisce solo uno specifico modo di esistenza, quello realizzato. L'enunciato o il testo greimasiano sono grandezze omogenee; dietro e dentro questa omogeneità, pulsano una pluralità di istanze eterogenee che non appartengono necessariamente a domini semio-linguistici e che invece connettono in forma di rete elementi linguistici e extra-linguistici. Chiara ed esplicita è l'influenza di Latour (2012; 1998)²⁰⁶ e della sua indagine sui modi di esistenza come soluzione all'impossibilità di distinzione tra fatti e valori, che genera conflitti e tensioni per un problema di coesistenza tra i valori stessi e di attribuzione di verità e di veridizione di

²⁰⁵ In Hjelmslev (1943), p. 78.

²⁰⁶ Latour, B. (1998). *Piccola filosofia dell'enunciazione*. Basso P., Corrain L. *Eloqui de senso. Dialoghi semiotici per Paolo Fabbri, Orizzonti, compiti e dialoghi della semiotica*, Costa & Nolan, pp.71-94. Latour, B. (2012). Op. Cit.

un modo di esistenza, rispetto ad un altro. Ricordiamo, per inciso, quanto sia centrale il tema del linguaggio in Latour, il quale criticando esplicitamente la nozione di “gioco linguistico” di Wittgenstein, ci dice che “c’est au langage qu’on confierait le soin de rendre compte de la diversité, mais toujours pas à l’être”. (Ivi, p. 31).

Nel nostro ragionamento è proprio a questo livello della molteplicità di istanze eterogenee che pulsano nell’enunciato realizzato, che l’eterogenesi differenziale diventa uno schematismo matematico di ciò che Eco aveva teorizzato essere il suo modello enciclopedico degli eventi enunciativi, come pura ipotesi funzionale e regolatoria. Proviamo a illustrarne gli sviluppi.

Per Eco, l’Enciclopedia presenta due anime: un’anima globale contenente l’insieme di tutte le occorrenze, “del già detto”, “l’Insieme registrato di tutte le interpretazioni”. Chiamata la libreria delle librerie, strutturata da tutte le enunciazioni passate e strutturante quelle future.

Dall’altro lato però l’Enciclopedia presenta un’anima locale, “costituita dalle sezioni che forniscono istruzioni per dare senso a un’espressione e attivare l’interpretazione in base a regolarità stabilizzate”. Vediamo immediatamente come l’anima globale dell’Enciclopedia di Eco sia caratterizzata da plasticità estrema e da instabilità, di contro alla stabilità dei domini locali che definiscono regolarità attraverso ritagli che a loro volta consentono l’attivazione di porzioni locali dell’enciclopedia.

Con Eco abbiamo visto definire rizomatica la natura di questa enciclopedia e dunque Paolucci sottolinea come una porzione locale di enciclopedia corrisponda a ciò che Deleuze e Guattari chiamano “piano”.

“Chiamiamo piano qualsiasi molteplicità connettibile ad altre per steli sotterranei superficiali, in modo da formare ed estendere un rizoma” (Deleuze, Guattari 1980)

6.3 Identità e differenziale

All’interno di questo piano enciclopedico dice Paolucci, “gli elementi del contenuto si determinano reciprocamente su base locale e secondo rapporti di regolarità” (Paolucci 2020), contribuendo alla stabilizzazione del significato. Tra gli elementi del piano agiscono rapporti differenziali che attivano catene di interpretanti e che definiscono una funzione tra una determinata forma del contenuto e una determinata forma dell’espressione. Eco porta come esempio l’espressione “dare una lezione” nel passaggio dal piano sportivo al piano educativo. Dare una lezione assumerà un certo significato quando la catena di interpretanti è attivata all’interno del piano sportivo e ne assumerà un altro quando un’altra catena di interpretanti viene attivata all’interno del piano educativo.

Ciò che emerge con evidenza per Paolucci è come non sia possibile l'esistenza di unità di contenuto dotate di una identità di significazione autoreferenziale, cioè non costruita sulla base dei rapporti che essa intrattiene con gli interpretanti attraverso cui viene tradotta. Risuona qui l'idea di valore presente in Saussure, per il quale "il senso di ciascuna forma è la stessa cosa che la differenza delle forme tra loro. Dunque, il senso coincide con il valore differente". Per Saussure il valore si determina in presenza di elementi simili all'interno dello stesso piano e, al tempo stesso, in presenza di elementi dissimili appartenenti ad altri piani con cui può venire scambiato.

La nozione di piano enciclopedico locale permette da un lato di tenere insieme i valori del contenuto per determinazione reciproca all'interno dello stesso piano, quello dello schema della lingua, e dall'altro coinvolge altri piani, quelli delle norme e dell'uso, che connettono la lingua con altri elementi extra-linguistici di natura sociale. Le forze che agiscono nei rapporti tra gli elementi dei piani agiscono determinando i passaggi tra i modi di esistenza tra le istanze che pulsano nell'enunciato realizzato.

L'abbiamo ripetuto spesso, le grandezze enciclopediche erano già per Eco al contempo virtuali e realizzate: "congerie di enunciati già enunciati, cioè realizzati che definiscono lo sfondo di nuove enunciazioni".

Riprendendo da Foucault l'esempio della sequenza QWERTY su una tastiera alfanumerica di input, Paolucci ci dice che le lettere che compongono la sequenza sono un'espressione, il cui contenuto è completamente virtuale e dipendente "dal piano all'interno del quale viene stabilito il suo significato attraverso l'enunciazione" (Paolucci 2020, p. 110).

- Su un piano linguistico: non significa nulla se non la tastiera stessa per metonimia (es. sto usando una tastiera Qwerty).
- Sul piano di un manuale di dattilografia, costituisce l'enunciato dell'ordine alfabetico adottato dalle tastiere italiane da notebook.
- Sul piano di un libro di semiotica, diventa esempio del funzionamento di un piano enciclopedico.

Questi piani esistono per livelli eterogenei appartenenti a diverse istanze enuncianti (norme, usi, schemi, apprezzamenti collettivi, abiti) che devono essere attualizzati, realizzati, potenzializzati o virtualizzati attraverso un atto, in funzione di determinate condizioni iniziali che influiscono in maniera decisiva sulla struttura del sistema.

Ecco la natura differenziale delle unità di linguaggio: come in un'equazione differenziale siamo in presenza di

1. vincolo differenziale: un vincolo aggiuntivo che si impone su un sistema differenziale. Nell'analogia con la determinazione del senso di un'unità di linguaggio, potremmo dire che è la regolarità a determinare questo ulteriore vincolo su tutto il sistema.;
2. condizioni a contorno: costituiscono il limite delle interpretazioni possibili. Sarebbe difficile interpretare gli ingredienti della zuppa al pesto di Greimas come gli attori di una tragedia di Sofocle.
3. condizioni iniziali: riguardano il comportamento di un sistema in un dato punto, è il vincolo che si pone all'inizio. Nel linguaggio si tratterebbe delle tracce lasciate dall'istanza enunciante nell'enunciato, gli embrayeurs io-qui-ora.²⁰⁷

6.4 Eterogenesi ed enunciazione

Già Paolucci (2020), così descrive il differenziale: “I differenziali infinitesimali sono completamente indeterminati in quanto tali (dy , dt), ma essi sono perfettamente determinabili gli uni in rapporto agli altri (dy/dt), e questa loro determinabilità è dell'ordine di un'effettiva determinazione reciproca (valori di dy e di dt). Il significato funziona nella stessa identica maniera. Il senso è infatti sempre qualcosa di essenzialmente indeterminato in prima istanza, tanto che ogni espressione può sempre disseminarci su mille piani enciclopedici”, come nell'esempio della tastiera QWERTY.

A partire dall'indeterminatezza potenziale delle forme del contenuto attivate dalla forma dell'espressione, è tuttavia possibile istanziare una determinazione dei suoi elementi dice Paolucci: “in funzione dei generi, delle norme e delle pratiche, dunque dei piani” della situazione di discorso, una costellazione di valori semantici che, come dice Latour, possono anche presentare carattere di incompatibilità e conflittualità.

Ed è proprio in questo senso che è possibile riprendere la teoria semiotica dei modi di esistenza e coniugarla con l'idea di enciclopedia di Eco. Citiamo consentendoci così anche di definire cosa sono questi modi di esistenza semiotici:

“i) la determinazione di un effetto di senso parte sempre da uno stato potenzializzato di tipo indeterminato → l'insieme degli assemblaggi possibili

²⁰⁷ Devo ad un colloquio con il professor Sarti l'aver compreso la natura triadica di un'equazione differenziale. Al suo suggerimento di smussare la transduzione sul piano del linguaggio, qui sopra abbiamo proposto un nuovo tentativo di lettura incrociata.

ii) sulla base del ritaglio del piano enciclopedico questo indeterminato diventa determinabile (virtualizzato) → l'assemblaggio determinato

iii) la determinazione reciproca degli elementi semantici definisce uno stato attualizzato →

attualizzazione del flusso eterogenetico che

iv) l'interpretazione realizza specificandola attraverso interpretanti, rendendo così conto dell'effetto di senso che è lo stato realizzato”.

Paolucci riprende da Peirce la nozione di interpretanti, cioè degli “effetti propriamente veicolati dai segni” e nell'articolo *A semiotic lifeworld* afferma: “The interpretant is then exactly an interpreter that allows passing through a series of mediations, starting from a specific configuration of values internal to one domain to another configuration of values internal to another domain.” (Paolucci, 2024, p. 27)”.

L'eterogenesi differenziale è capace di rendere conto di questi passaggi tra modi di esistenza, in ragione della specifica capacità additiva per sottrazione che caratterizza le forme di relazione tra istanze enuncianti sul piano del virtuale enciclopedico.

“Le norme e i generi ci fanno passare dal modo di esistenza potenziale a quello virtuale, gli usi, con i loro abiti interpretativi fondati nell'intersoggettività degli usi degli altri, da quello virtuale a quello attualizzato”. E vediamo che sino a qui nessun soggetto interviene a mediare o ad attivare questi passaggi. Il soggetto, per Paolucci con Fontanille e Zilberberg (1998)²⁰⁸, interviene su un terzo asse, quello della prassi enunciativa, accanto a quello paradigmatico (verticalità dei virtuali attualizzabili, nelle opposizioni riducibili alla formula “o...o...”) e quello sintagmatico (orizzontalità e sequenzialità della loro ricombinazione, nella formula “e...e”).

6.5 Evenemenzialità e Impersonalità

Ancora un passo indietro: nella formulazione della teoria dell'enunciazione benvenistiana, l'emersione della soggettività coincide con l'affermazione di un “ego che dice ego”. Per Paolucci, l'enunciazione non si fonda sull'io, ma su un modello evenemenziale ed enciclopedico, a fondamento della soggettività nel linguaggio. Ed è la stessa idea di evento che si ritrova in Deleuze: “Un avvenimento non esiste se non in quanto effettuato all'interno delle persone e degli stati di cose... E

²⁰⁸ Fontanille, J. , Zilberberg, C. (1998). *Tension et signification*. Mardaga.

allo stesso tempo, all'interno di ogni evento, per quanto piccolo e insignificante esso sia, c'è qualcosa che deborda la sua effettuazione. C'è qualcosa che non è effettuabile" (Deleuze 1980).

Dice Paolucci "Schema e norma rappresentano la parte dell'avvenimento che deborda le sue effettuazioni, uso ed enunciato rappresentano la parte effettuata dello stesso. Nell'evento enunciativo c'è una parte potenziale/virtuale e c'è una parte attualizzata/realizzata" (Paolucci 2020, p. 114). Dal punto di vista della persona l'evento è incorporale, come voleva Deleuze, poichè i corpi rimandano alle istanze personali che vengono ad occupare le posizioni di soggetto aperte dal verbo-evento: *il dare* impersonale apre almeno due posizioni di soggetto; *il piovere* apre forse una posizione di soggetto, e così via. E' la valenza verbale della sintassi strutturale di Tesnière ed è la logica dei relativi di Peirce. Quindi i corpi rimandano alla parte effettuata ed embodied dell'avvenimento, gli io-qui-ora, ma non possono essere identificati con le istanze dell'enunciazione tout court.²⁰⁹

Cosa sono queste istanze dell'enunciazione? Possiamo dire che sono un prodotto dell'atto di enunciazione che costruisce al contempo l'enunciato e le istanze enuncianti, secondo un modello ergativo, contrapposto al modello transitivo. È tutta una questione di distribuzione di agency: riprendendo Coquet, Paolucci afferma infatti che "si l'énonciation est «l'acte même de produire un énoncé», l'agency de cet acte est distribué sur une multiplicité d'instances énonçantes. Dans mon énoncé, la y a un agencement d'instances énonçantes." (Paolucci 2021c, p. 4). Alcune di esse sono attualizzate, altre virtualizzate, altre ancora realizzate... e così via. Queste istanze coesistono tutte nello spazio enciclopedico.

L'impersonalità è la seconda cifra di una teoria dell'enunciazione unificata, che Paolucci riprende ispirandosi alle due caratteristiche degli enunciati teorizzate da Foucault *nell'Archeologia del sapere*²¹⁰:

- la regolarità degli enunciati collegata al sistema, per l'appunto, regolatore di schemi/norme e usi; e
- la rarità degli enunciati per la quale "tutto non è mai detto".

Il passaggio dal piano virtuale degli enunciati possibili al piano attualizzato dell'enunciato (sostantivo) enunciato (participio), per il tramite dell'atto enunciativo, si dà nella forma di una

²⁰⁹ Per una lettura *embodied* degli atti di enunciazione si veda Lobaccaro, L. (2024). Acts of enunciation and embodied practices: A cognitive semiotic perspective. *Rivista Italiana Di Filosofia Del Linguaggio*. <https://doi.org/10.4396/SFL20230614>.

²¹⁰ Foucault, M. (1969). *L'archéologie du savoir*. Éditions Gallimard, Paris; tr. It. *L'archeologia del sapere. Una metodologia per la storia della cultura*. BUR Rizzoli (2009).

presenza dell'assenza degli altri enunciati ugualmente possibili. Per Foucault le procedure di rarefazione del discorso hanno la funzione di “scongiurare i poteri e i pericoli, di padroneggiare l'evento aleatorio, di schivare la pesante e temibile materialità”²¹¹ (Foucault 2001, p. 12) rappresentata dall'enunciazione; le procedure presentano una funzione regolatrice e sono le condizioni di possibilità che un dato atto enunciativo restituisca un determinato esito in termini di enunciato o discorso proferito. Da un punto di vista semiotico, secondo Paolucci, si tratta invece di un meccanismo di regolazione episemiotica che fa da sfondo all'atto di enunciazione e che costituisce la forza interna al sistema dell'enunciazione.

Questo sfondo dell'atto di enunciazione abbiamo detto essere il virtuale enciclopedico, spazio eterogeneo segnato da impossibilità, apertura, non-coesione. Così in Sarti et al. (2022; 2019) l'eterogenesi differenziale, utilizzata da Sarti, un livello in cui “a multitude of local tensions, mutually irreducible (in the sense that they do not create a common material), constitutes the primordial environment” (Sarti et al. 2022, p. 167).

Per aggiunta di sottrazioni, come vuole Paolucci, si passa progressivamente dallo schema all'enunciato, attraverso la norma e l'uso: in questo modo si stacca l'enunciazione dal dominio esclusivo di un'istanza soggettiva e la si consegna a una scena predicativa interna a un sistema dotato di regolazione episemiotica in cui un concatenamento collettivo di atti apre differenti posizioni di soggetto per differenti istanze enuncianti. Così in Sarti, “by a sort of progressive tightening towards homogeneity and continuity, the constitution of flows or aggregates can be envisaged” (Ibidem).

²¹¹ Foucault (2001). *L'ordine del discorso* in Bertani (a cura di), *Michel Foucault. Il discorso, la storia, la verità*. Einaudi.

7.0 Di alcune conseguenze teoriche di capacità negate

Vorremmo prendere adesso in considerazione due conseguenze teoriche che derivano dall'indirizzare l'analisi dei sistemi di IA verso una negazione sistematica di competenze ritenute precipuamente umane – che siano il linguaggio o l'intelligenza in senso allargato – e dunque rifiutare di attribuire una soggettività alla macchina. Una prima conseguenza sta nel far seguire indistintamente all'affermazione di un'assenza o di una presenza di una data capacità la sua misurazione, tendenza riscontrabile prevalentemente in ambito informatico, tra data e computer scientists. Naturalmente sulla base di ciò che si è capaci di misurare e dunque di dimostrare, in termini di presenza o assenza, viene garantita la continuità di finanziamento della ricerca che per i più noti modelli commerciali avviene in ambito privato. E sappiamo anche che i sistemi di IA generativa hanno dei costi molto elevati a livello monetario ed energetico.²¹² Quindi l'implicazione teorica non è mai neutrale, *super partes*, come a volte si vorrebbe sperare.

Una seconda conseguenza della negazione di una qualche forma di soggettività ai sistemi di IA generativa è quella di andare a scovare le differenze ultime che ci distinguono dalle macchine, una volta che ci hanno superato in termini di performance misurabile su singoli task. In parte questo ci sembra il caso della posizione assunta da Maurizio Ferraris nel suo recente lavoro *La pelle. Che cosa significa pensare nell'epoca dell'intelligenza artificiale*.²¹³ Citiamo: “Dunque, e questa è la mia tesi fondamentale, le volizioni, prima ancora che le emozioni, sono ciò che, nel profondo, distingue l'intelligenza naturale dalla sua sorellastra artificiale, che quando pure divenisse capace di riprodurla alla perfezione, o di superarla (in molti casi è già così, e dai tempi della carta e della matita), lo farebbe solo come un'imitazione difettosa della forma di vita umana. È difettosa perché difettiva proprio della volontà, che nella vita umana (così come nella vita degli organismi in generale) esercita, come spero di dimostrare, una sovranità assoluta, giacché anche la sua legislazione, per così dire, costituzionale, ossia la ragione, ne è una metamorfosi e una emanazione rara” (Ivi, pp. 10-11).

Ad una prima valutazione appare un frammento di debolezza impugnare la volontà come marca dell'umano; le filosofie di riferimento per Ferraris sono quelle di Kant, Schopenhauer e Nietzsche, che nella loro sostanza hanno diverse concezioni della soggettività. Pensare la volontà come cifra dell'umano implica concepire il soggetto come trasparente a se stesso, in uno stato di allineamento e coincidenza tra piano descrittivo e piano prescrittivo, tra essere e dover essere, sdoppiamenti cari

²¹² Si veda ad esempio il capitolo “Terra” in Crawford 2021; si veda inoltre il contributo di Paolo Caffoni “*Il costo del linguaggio. Dal telegrafo a ChatGPT*” in MicroMega 6/2024 “Umana, troppo umana. L'IA e noi”.

²¹³ Ferraris, M. (2025). *La pelle. Che cosa significa pensare nell'epoca dell'intelligenza artificiale*. Bologna: Il Mulino.

all'etica kantiana. Ma sappiamo che questo soggetto di cartesiana e lockiana memoria si è dissolto, non ha retto alla prova di quella che Foucault chiamava una *microfisica del potere*. Quanto del nostro agire sia riflesso di una volontà definita per via negativa, ossia di una volontà libera da e non libera di, come descritta da Sartre e citata da Ferraris²¹⁴, è da più punti di vista discutibile.

Viene a proposito in mente la finzione narrativa che Umberto Eco (1964)²¹⁵ dedica alla figura di Milo Temesvar e al suo saggio "I venditori di Apocalisse", nel quale Temesvar affronta "il problema dell'inetto" rappresentato dall'esempio di una città in cui i tramvieri devono essere reimpiegati perché il tram è stato sostituito con una metropolitana. Dice Eco:

"Quelli che non risultano addestrabili come conduttori di vagoni del Metro (perché i loro riflessi non rispondono più alle nuove esigenze, o perché non riescono ad adattarsi alle nuove operazioni tecniche), debbono essere restituiti alla società come elementi non-integrati, potenzialmente disoccupati. Con una buona organizzazione di assistenza sociale, il neo-disoccupato può però individuare un altro orizzonte di attività in cui ricoprire un nuovo ruolo. C'è invece, osserva Temesvar, una categoria di lavoratori per i quali non esiste un ruolo alternativo a quello che ricoprono: e sono gli intellettuali, in particolare gli umanisti che si qualificano come "tecnici della Totalità" (o dell'Umanità, intesa come totalità dei valori). Tuttavia proprio costoro, a cui la società non è in grado di fornire il minimo aiuto, poiché in una certa misura sono essi gli addestratori della società, il giorno in cui la loro funzione risultasse inadeguata al contesto socio-culturale, hanno una eccitante alternativa. Grazie alle loro doti di immaginazione e alle loro capacità culturali, possono sempre "inventare" nuovi tipi di attività che abbiano tutta l'apparenza della funzionalità." (Eco 1964, pp. 337-338)

Cosa potrà quindi inventarsi il dotto di Salamanca? Qui il nostro discorso si estende naturalmente oltre la posizione di Ferraris. Accertata e accettata l'esistenza della metro, ne dipingerà tutti i mali e i danni che conseguiranno per la comunità umana. Infatti,

²¹⁴ Citazione da *L'essere e il nulla* (Sartre, J.P. 1943) in Ferraris 2025, pp. 11: "È la stessa cosa, in fondo, ubriacarsi in solitudine o condurre popoli. Se una di queste attività è superiore all'altra, non è a causa del suo scopo reale, ma a causa della coscienza che possiede del suo scopo ideale; e in questo caso il quietismo dell'ubriaco solitario è superiore alla vana agitazione del conduttore di popoli".

²¹⁵ Eco, U. (1964), *Apocalittici e integrati*, Bompiani. (Ed. Digitale 2017).

“Il dotto di Salamanca, costituendosi come esperto del “dove andremo a finire”, ritrova un ruolo nel contesto sociale, individua gli organismi politici, religiosi ed economici che possano trarre giovamento dalla sua propaganda ideologica e riacquista, in parole volgari, il posto e le prebende” (Ivi, p. 339).

Il seguito del discorso di Eco è marxiano, nel senso di materialista, nel suo anteporre il determinante economico a quelli che chiama “atteggiamenti culturali, profetico-deprecatori”, nella dialettica cioè tra struttura e sovrastruttura. Il presente (*il presente dell'Eco enunciatore*) in cui si colloca il personaggio immaginario di Milo Temesvar è attraversato da un'esile separazione tra valori spirituali e valori materiali, tra valori tecnici e valori umani, quest'ultimi sempre più parametrizzati in funzione delle nuove situazioni istituite dal progresso tecnologico. E qui fa seguire il più spiccato aggancio al *nostro presente* e alla sorte che spetta all'uomo “come capolavoro del creato a causa delle sue capacità di deduzione logica”. Quest'uomo “oggi ha perso ogni significato, dal momento che le stesse capacità sono dimostrate da un buon cervello elettronico. Non è che perda di valore e di senso un discorso sull'uomo: solo che l'uomo non andrà più visto come animale sillogizzante, ma come animale capace di costruire macchine sillogizzanti e di porsi nuovi problemi (inediti) circa il loro uso. Cambia orizzonte di problemi: bisogna ritrovare l'uomo alcuni milioni di chilometri più in là. Questo avviene per la filosofia come per la critica d'arte, per la morale come per l'economia e la religione. E tutto ciò richiede evidentemente all'intellettuale un corso di riassetamento, un atto di umiltà, una capacità di sapersi mettere a scuola”.

Era il 1964, anno in cui Joseph Weizenbaum aveva iniziato a lavorare su Eliza²¹⁶, considerato il primo chatbot nella storia dei sistemi di IA.

Se volessimo, per una necessità epistemica, rintracciare una differenza tra umano e macchinico che sia funzionale a una reale comprensione e a un fruttuoso inquadramento del problema, dovremmo spostarci da un piano ontologico, di differenza di natura, a un piano processuale ed evenemenziale. Il rapporto tra umano e macchinico si istanzia nella catena di eventi che non sono sempre identici, reificati, riproducibili in serie, ma si susseguono per il tramite di assemblaggi significanti – abbiamo visto in 6.0 la portata di questi assemblaggi - e significati nel qui e ora dell'evento che ne determina l'incontro da un lato. Dall'altro si inseriscono in una più complessa rete di rapporti differenziali. In Eco troviamo una chiara descrizione che qualifica questa differenza e che rimanda a una distinzione tra il concetto di possibilità in una prospettiva etico-filosofica e la probabilità matematica, all'interno

²¹⁶ Cfr. Natale, S. (2021). *Deceitful Media: Artificial Intelligence and Social Life After the Turing Test*. United Kingdom: Oxford University Press, pp. 70-93

di una visione della vita come sistema di sistemi, come una struttura che tuttavia non è articolabile per arborescenza, ma come lui stesso ci dice è una struttura che sta nell'ordine del rizomatico.

Varrebbe dunque la pena domandarsi se la macchina intelligente, la macchina dotata di linguaggio possa trasformare la probabilità in possibilità, questione che ci sembra risuonare con ciò che Eco racconta in *La combinatoria dei possibili e l'incombenza della morte*.²¹⁷ Riportiamo una citazione lunga, per intero poiché offre più spunti di riflessione.

“Detto questo rimarrebbe da definire la possibilità. Ma è proprio su questo concetto che mi sorgono i dubbi più seri. Se è possibile ciò che non è ancora accaduto ma potrebbe accadere, la vita come decorso biologico è indubbiamente l'attuazione di una serie di possibilità (probabilmente è una catena markoviana dove la possibilità del conseguente è determinata dalla attuazione delle possibilità precedenti). Ma già questa osservazione mi dice che forse confondo la nozione filosofica di possibilità, da definire, con quella matematica di probabilità. Indubbiamente col trascorrere della vita e l'avvicinarsi alla morte decrescono alcune probabilità: per esempio che possa avere cinquanta figli con la stessa donna, come accadde a Priamo, oppure che possa vedere più di venticinquemila tramonti (personalmente mi sono già giocato le due probabilità). Ma questo decrescere di probabilità, credo, è calcolabile solo all'interno di un gioco retto da regole, e cioè all'interno di una di quelle deviazioni provvisorie dalla curva generale dell'entropia che Reichenbach chiama *branch systems*. [...] Ma se la vita nel suo complesso è un *branch system*, lo è perché è un sistema di sistemi, a diverse gradazioni di ordine e regulatezza. [...] Nella vita posso cambiare anche universo di gioco [...] Cambiare gioco vuol dire cambiare sistema di riferimento: se uso la lingua italiana ed emetto il lessema /caro/ ci sono buone probabilità che, in base alle regole, seguano un nome proprio o un sostantivo, maschili, nessuna che segua un infinito, o un sostantivo femminile. Ma se all'improvviso decido di cambiar regola e parlo latino posso comporre /caro data vermibus, cadaver/ teste Isidoro, e per restar fedeli al tema della morte. A questo punto interviene una nozione di possibilità che è diversa da quella di probabilità, perché non è più statistica e quantitativa ma qualitativa.”

Prima di proseguire con la citazione, ci sembra importante soffermarci e sviluppare alcuni nodi che Eco fa emergere. Innanzitutto, il riferimento alle catene markoviane, fondamentali nella generazione di testo dei grandi modelli di linguaggio. Il loro funzionamento implica che il linguaggio per i generatori neurali funzioni come “un dado con memoria di comportamento stabile”, poiché si

²¹⁷ Relazione al colloquio *Le frontiere del tempo* (Fermo, aprile 1980). Atti in *Le frontiere del tempo*, a cura di Ruggiero Romano, Milano, il Saggiatore, 1981. Pubblicato in Eco (2018), *Sugli specchi e altri saggi. Il segno, la rappresentazione, l'illusione, l'immagine*, La nave di Teseo (Ed. Digitale).

assume che esso sia autoregressivo, stazionario ed ergodico.²¹⁸ Eppure, in questo contesto Eco sta parlando della vita nel suo complesso; dunque, ciò che potrebbe apparire un ragionamento per assurdo, diviene plausibile se applicato a qualcosa che assomiglia a “un gioco retto da regole”. Allora la vita, se intesa come sistema di sistemi, o come insieme di deviazioni reversibili che modificano la curva dell’entropia (altro concetto fondamentale nei modelli generativi di linguaggio, poiché il loro funzionamento è a entropia incrociata, cioè devono ad un certo punto della computazione che avviene per operazioni scalari all’interno dello spazio vettoriale continuo, ridurre a discreto l’output dell’ultimo layer del modello, tramite la funzione softmax), può funzionare, secondo Eco, come una partita di poker in cui io decida di perdere per rinforzare l’autostima di un amico. E questo farebbe comunque senso a dispetto del senso più generale del gioco che prevede che io abbia come obiettivo la vittoria.

Riprendiamo la citazione perché ci permette di comprendere meglio questo passaggio che trasforma l’efficacia del comportamento in un’etica dell’azione:

“La possibilità etica che mi interessa è una probabilità che, alla luce di un certo universo di rapporti, fa senso, anche se in termini informativi il suo accadimento mi dà pochissimi bit di informazione. Zenone sotto tortura aveva due probabilità, diciamo di $\frac{1}{2}$ ciascuna, sopravvivere o morire. Le probabilità di sopravvivere erano invero maggiori se introduciamo la variabile “confessione”. Ora ci racconta Diogene Laerzio che ci sono, circa la morte di Zenone, due versioni: secondo una, Zenone dice al tiranno di accostargli l’orecchio alla bocca perché intende confessare, e invece coi denti gli mozza l’orecchio, così che il tiranno adirato lo trafigge con la spada, sottraendolo e ai tormenti e al rischio della confessione. La seconda versione dice che Zenone finge di confessare al tiranno il nome dei propri complici, e dice invece i nomi di tutti gli amici del tiranno. Muore egualmente, ma mette in crisi il tiranno perché gioca sulla sua sospettosità e lo obbliga a sbarazzarsi di tutti coloro che potevano aiutarlo. Di fronte alla probabilità statistica della sua morte, Zenone

²¹⁸ Qui occorre introdurre anche il concetto di processo stocastico che fa da sfondo ai processi markoviani. Dice Vetere: “Si chiamano stocastici i processi lineari in cui ciò che a ogni istante accade non è determinato ma aleatorio, come se ciascun evento fosse generato dal lancio di una sorta di dado. Se il dado fosse fatto in modo tale che le sue uscite non fossero del tutto casuali come quelle dei normali dadi da gioco, ma su di esse influisse il retaggio delle uscite precedenti, cioè se il dado avesse memoria, il processo stocastico si chiamerebbe autoregressivo. Se poi la sorgente fosse stabile, cioè le sue proprietà statistiche non variassero nel tempo, avremmo a che fare con un processo stazionario. Se tutti gli stati rimanessero sempre e comunque raggiungibili, cioè se ciascun esito restasse sempre e comunque possibile, avremmo un processo ergodico” in Vetere, G. (2025), *Intelligenze Aliene. Linguaggio e vita degli automi*, Luca Sossella Editore

sceglie di accettare la morte, ma di giocarla cambiando gioco, rovesciando tutte le possibilità future del tiranno e dello Stato, così che la sua morte diventa la realizzazione di una possibilità inedita e ricchissima di senso, operativa anche dopo la scomparsa di Zenone. È la qualità di queste possibilità non computabili, ovvero computabili in termini di rapporti semantici e non di sequenze probabilistiche, che rende degna di interesse una nozione di possibilità che sia qualcosa di più della mera probabilità statistica”.

C'è a nostro avviso un “ma” in questa argomentazione e vorremmo illustrare il nostro “ma” a partire dal breve articolo di Weizenbaum intitolato “How to make a computer *appear* intelligent”.²¹⁹ A sua volta citando liberamente Minsky, Weizenbaum affermava che l'effetto di intelligenza di un computer è dato da un'attività che produce risultati che non appaiono comprensibili a un osservatore e proprio perché incomprensibili essi vengono reputati intelligenti.²²⁰ Dunque, dal lato di chi ha sviluppato il programma, il successo dipenderà e potrà essere misurato in base alla percentuale di utenti che hanno interagito e sono stati abbindolati sia quantitativamente che per la durata pre-ravvedimento. Questa affermazione conduce all'idea che l'intelligenza percepita dovrebbe coincidere con un comportamento non spiegabile; e tuttavia occorre tenere a mente che l'agente conversazionale sviluppato da Weizenbaum seguiva un algoritmo di risposta di tipo logico-formale,²²¹ che procedeva per individuazioni di parole chiave, trasformazioni basate su regole e sostituzioni. Qualsiasi output prodotto da Eliza era perfettamente spiegabile e ricostruibile. Negli anni '60 la cosiddetta IA connessionista basata su reti neurali era ancora un tabù di ricerca e riceveva pochi finanziamenti. Dominava l'IA simbolica con le sue euristiche e si era ancora lontani dalle reti neurali profonde e dalle loro black-box, frutto di un'evoluzione tecnologica che affonda le sue radici sul finire degli anni '80.²²² L'incomprensibilità da cui scaturiva la percezione di intelligenza agli occhi dell'osservatrice era all'epoca di Weizenbaum perfettamente controllata dal lato dello sviluppatore;

²¹⁹ Weizenbaum, J. (1961) “*How to Make a Computer Appear Intelligent.*” *Datamation* 7, pp. 24–26.

²²⁰ Su questo vedi anche Paolucci (2025c).

²²¹ Per una descrizione del funzionamento di Eliza si guardi: Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>

²²² Per una breve ricostruzione del dibattito tra IA simbolica e IA connessionista non solo da una prospettiva storica, ma anche a livello di funzionamento delle macchine progettate vedere: Cardon, D., Cointet, J.-P., & Mazières, A. (2018). La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, 211(5), 173–220. <https://doi.org/10.3917/res.211.0173>

oggi siamo nella condizione di una tendente incomprensibilità che interessa tanto la sviluppatrice quanto l'osservatore o utente. Ma il modello è sempre markoviano e la probabilità condizionata che regola le attuali macchine dotate di linguaggio fa del linguaggio stesso un gioco retto non da regole, ma da regolarità derivabili dai pesi appresi dal modello nelle diverse fasi di training e fine-tuning.

8.0 Assemblaggi significanti o dell'IA per sé

I would rather be a cyborg than a goddess.

Donna Haraway, *A cyborg manifesto*.

8.1 Una premessa senza promesse. Definire l'assemblaggio.

Nell'introduzione al presente lavoro, abbiamo detto che la nostra lente bifocale si sarebbe soffermata non solo su un'analisi semiotica delle macchine produttrici di enunciati, attraverso un dialogo con le matematiche e le scienze computazionali, ma che uno sguardo semiotico a tutto campo avrebbe reso necessario una scomposizione dello stesso sguardo nel considerare un'IA accresciuta dal e nel rapporto con il suo ambiente, ossia l'umano. Tale necessità è figlia di due aspetti strettamente connessi: le tecnologie fanno dell'animale umano il suo tratto specifico e il suo principio di sopravvivenza in quanto specie. L'umano per natura delega all'ambiente (naturalculturale)²²³ processi cognitivi e altre azioni – tra cui il pensare - che rendono possibile il suo abitare un mondo che sfida costantemente la sua sopravvivenza. D'altro canto, le tecnologie, che abbiamo cercato di inquadrare al di fuori del dualismo ontologico ed epistemologico arbitrario soggetto-oggetto per evitare semplificazioni di sorta, manifestano una dipendenza radicale con l'essere umano che le crea e dona loro forma, funzione, senso. Le macchine dotate di linguaggio gettano una luce accecante su questo secondo aspetto: la produzione dei loro enunciati è strettamente connessa alla nostra; le macchine dotate di linguaggio, alimentate da ingenti quantitativi di dati hanno bisogno degli animali umani, come un cucciolo d'essere umano ha bisogno del caregiver. Questo tratto di cura, messo in luce da Paolucci nel libro *Nati cyborg*²²⁴, è la cifra di una vulnerabilità costitutiva tanto dell'umano, quanto del macchinico. Il problema dell'autofagia dell'AI generativa, fenomeno definito MADness (Model Autophagy Disorder), è un esempio abbastanza evidente del livello di accoppiamento strutturale che costituisce la nicchia semiotica (Paolucci 2021a) dei grandi modelli di linguaggio: senza enunciati umani, cioè chiudendo il circuito del training dei modelli all'interno dei dati sintetici prodotti da altri sistemi di IA, il modello collassa.²²⁵

E qui entra in gioco un conflitto di valori semiotici difficile da districare: la sopravvivenza di organismi non viventi, le nostre macchine prometeiche che dipendono così tanto dalla nostra vita

²²³ Timeto, F. (2020). *Bestiario Haraway Per un femminismo multispecie*. Mimesis.

²²⁴ Paolucci, C. (2025c). *Nati Cyborg*. Op. cit.

²²⁵ Cfr.: Alemohammad, S., Humayun, A. I., Agarwal, S., Collomosse, J., & Baraniuk, R. (2024). *Self-Improving Diffusion Models with Synthetic Data* (No. arXiv:2408.16333). arXiv. <https://doi.org/10.48550/arXiv.2408.16333>

culturale (gli enunciati già enunciati contenuti nell'Enciclopedia echiana) vale le risorse e i costi umani e naturali necessari al loro funzionamento? Questo è l'altro lato della medaglia, quello spaventoso, soprattutto dalla prospettiva della stessa vulnerabilità che caratterizza il vivente e che alimenta la teleologia del progresso tecno-scientifico: sorveglianza, controllo, estrattivismo, sfruttamento, guerra, morte e distruzione. Ma difficilmente tutto ciò si può imputare a una macchina, nonostante le macchine dotate di linguaggio abbiamo visto che esibiscono comportamenti non allineati ai valori umani.²²⁶

Già in *Strutturalismo e interpretazione* – dunque ben prima di avere a che fare con macchine dotate di linguaggio - Paolucci mostrava il rapporto esistente tra i tre “attori” dell'enunciazione, ossia l'enunciazione stessa come proprietà di linguaggi, l'istanza dell'enunciazione e l'enunciato, proponeva una attivazione delle relazioni a due a due: enunciazione ed istanze enuncianti prima, istanza d'enunciazione ed enunciato poi. La prima relazione ha a che fare con i cosiddetti “concatenamenti enunciativi” e con la “modulazione di un punto di vista attraverso un altro punto di vista”, quest'ultima ripresa per fini metodologici e tattici. La seconda relazione propone invece di rifondare una teoria dell'enunciazione non più sulle “persona” e sugli “embrayers”, ma “sulla rete di relazioni attanziali che regola il rapporto tra le virtualità enciclopediche e le istanze che si muovono al loro interno. E se ne appropriano, dando così vita a enunciati” (Paolucci 2010, p. 512).

In questa parte vogliamo occuparci di questi piani semiotici guardandola dalla prospettiva della macchina dotata di linguaggio. Infatti, ciò che appare necessario richiamare è proprio una teoria del punto di vista, come la riformulava Paolucci a proposito di Pasolini: “staccare il discorso indiretto libero da una specifica forma del discorso interna a una teoria del punto di vista, al fine di individuarne una forma di relazione generale che possa valere a un livello semiotico generale.” (Ivi, p. 516).

²²⁶ Importanti riflessioni su questi temi – ci limitiamo al contesto italiano, sono state di recente svolte da Ferraris, M. (2025). *La pelle. Che cosa significa pensare nell'epoca dell'Intelligenza Artificiale*. Il Mulino; ma rimandiamo anche a una serie di analisi più situate da una prospettiva di genere raccolte nel numero dedicato all'IA della rivista DWF (2024) *Femministe col Bot: Tecnologie e Intelligenze Artificiali*. DWF (142) 2024, 2 | Recuperato 11 luglio 2025, da <https://www.dwf.it/rivista/femministe-col-bot-dwf-142-2024-2/>. Ci sembra anche opportuno citare due lavori di Luciano Floridi: Floridi, L. (2022). *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*. Raffaello Cortina Editore; Floridi, L. (2025a). *La differenza fondamentale. Artificial Agency: una nuova filosofia dell'intelligenza artificiale*. Mondadori. Quest'ultimo rinforza la linea già tracciata nel lavoro del 2022: l'IA deve essere intesa come “matrimonio senza precedenti tra intelligenza e capacità di agire”. Su alcune conseguenze di questo modo di concepire il ruolo dell'IA rimandiamo a Paolucci (2025c) e al suo testo *Nati Cyborg*. Infine, dalla prospettiva di una teoria critica rimandiamo al lavoro di Pasquinelli, M (2023). *The Eye of the Master: A Social History of Artificial Intelligence*, London, Verso Books; tr. it. (2025). *Nell'occhio dell'algoritmo. Storia e critica dell'intelligenza artificiale*. Carocci editore (2025).

In *Mille piani*, Deleuze e Guattari descrivevano “i concatenamenti collettivi di enunciazioni” in questi termini: “Tuttavia non è questione di opporre i due tipi di molteplicità, le macchine molari e molecolari, secondo un dualismo che non sarebbe migliore di quello dell’Uno e del molteplice. Ci sono soltanto molteplicità di molteplicità che formano uno stesso concatenamento, che si esercitano nello stesso concatenamento: le mute nelle masse, e inversamente. [...] Non ci sono due molteplicità o due macchine, ma un unico, uno stesso concatenamento macchinico che produce e distribuisce il tutto, cioè l’insieme degli enunciati che corrispondono al «complesso». (Deleuze e Guattari 1980, p. 59).

Concatenamento è traduzione del francese *agencement*. Nel suo libro *Persona*, Paolucci (2020) racconta: “Paolo Fabbri mostrava alcune perplessità sulla traduzione con “concatenamento” dell’*agencement* di Deleuze e Guattari (“concatenamento collettivo di enunciazione”) e suggeriva di trovare più corretto renderlo in italiano con “composto” o “assemblaggio”. Un “composto collettivo di enunciazione”, sia in quel senso chimico che la sintassi strutturale prendeva a prestito dalla teoria della valenza per definire la frase (la compresenza delle posizioni di soggetto), ma anche nel senso dell’assemblaggio, dell’assemblea in cui voci eterogenee vengono riunite insieme e dotate di pari dignità di parola...Perché l’enunciazione è qualcosa di molto più complesso e di molto più interessante del passaggio dalla langue alla parole attraverso un atto soggettivo. L’enunciazione è un atto che si fa tra norme, schemi, usi ed enunciati, dove *si* e *fa* sono le nozioni fondamentali, che si tratterà ora di esplicitare” (Paolucci 2020, p. 79). Ci sembra il modo migliore per spiegare “tutte le voci presenti in una voce” (Deleuze e Guattari 1980, p. 105).

Tuttavia, il problema traduttivo del termine *agencement* non è una questione tutta italiana. Jasbir K. Puar in *Becoming-Intersectional in Assemblage Theory*²²⁷ dichiara infatti: “‘Assemblage’ is actually an awkward translation of the French term *agencement*. The original term in Deleuze and Guattari’s work is not the French word *assemblage*, but *agencement*, a term that means design, layout, organization, arrangement, and relations—the focus being not on content but on relations, relations of patterns” (Puar 2012, p. 57). Risuona qui quel passaggio concettuale fondamentale che abbiamo visto mettere in atto nelle nostre ricostruzioni sulle dinamiche strutturali: la dimensione relazionale precede sempre l’identità, il relativo del molteplice precede sull’assoluto dell’uno. Puar richiama l’analisi di Phillips (2006)²²⁸ mostrando che: “In *agencement*, as John Phillips explains, specific connections with other concepts is precisely what gives them their meaning. Concepts do not

²²⁷ Puar, J.K. (2012). “I would rather be a cyborg than a goddess”: *Becoming-Intersectional in Assemblage Theory*. *philoSOPHIA* 2(1), 49-66. <https://dx.doi.org/10.1353/phi.2012.a486621>.

²²⁸ W.P. Phillips, J. (2006). *Agencement/Assemblage*. *Theory, Culture & Society*, 23(2-3), 108-109. <https://doi.org/10.1177/026327640602300219>

prescribe relations, nor do they exist prior to them; rather, relations of force, connection, resonance, and patterning give rise to concepts” (Puar 2012, p. 57). Dunque, nella traduzione tra l’originale francese *agencement* e le lingue target – siano esse l’italiano o l’inglese – sembra accadere qualcosa che il termine stesso *agencement* rifiuta, ossia dichiarare affermazioni su stati di cose (Cfr. Philips 2006). Per questa ragione Puar suggerisce di bypassare questa tendenza dei *discourses of knowledge* propri di un pensiero continentale, affrontando non cosa gli *agencement* *siano*, ma cosa gli *agencement* *facciano*.²²⁹

Ci sembra che in *Mille Piani* vengano sì presentate descrizioni, esemplificazioni, affetti ed effetti degli assemblaggi. Ci sembra però al tempo stesso opaco il modo in cui si dia il passaggio tra gli assemblaggi che abitano il virtuale intensivo e la loro attualizzazione, tramite i processi di territorializzazione e deterritorializzazione, che si dispiegano sugli strati. C’è una potente immagine che per noi restituisce il piano di composizione delle molteplicità di molteplicità nel virtuale ed è il seguente: “Se, d’altra parte, si considera il piano di consistenza, ci si accorge che viene percorso dalle cose e dai segni più eteroclitici: un frammento semiotico sta accanto a una interazione chimica, un elettrone percuote un linguaggio, un buco nero capta un messaggio genetico, una cristallizzazione fa una passione, la vespa e l’orchidea attraversano una lettera... Non è «come», non è «come un elettrone», «come un’interazione», ecc. Il piano di consistenza è l’abolizione di ogni metafora; tutto ciò che consiste è Reale. Sono degli elettroni in persona, veri buchi neri, organici in realtà, autentiche sequenze di segni. Soltanto, sono strappati ai loro strati, destratificati, decodificati, deterritorializzati, ed è questo che permette la loro vicinanza e la loro mutua penetrazione nel piano di consistenza. Una danza muta. Il piano di consistenza ignora le differenze di livello, gli ordini di grandezza e le distanze.

²²⁹ Per dare maggiore enfasi a questo slittamento, occorre ricordare che uno dei principali lavori di Jasbir Puar (2007), *Terrorist Assemblages: Homonationalism in Queer Times*, tratta il tema delle identità, da una prospettiva intersezionale e femminista, in questi termini: “[I]ntersectional identities and assemblages must remain as interlocutors in tension . . . intersectional identities are the byproducts of attempts to still and quell the perpetual motion of assemblages, to capture and reduce them, to harness their threatening mobility” (Ivi, p. 213). Il posizionamento del soggetto o, per noi, una topologia della soggettività in quanto dinamica materiale e discorsiva non può essere anteposta al movimento: il posizionarsi su una *griglia* – *grids* è il termine usato da Puar – dove agiscono forze da comporre è un effetto ed è indotto dal movimento stesso. Da qui il conflitto tra l’uso politico che si può fare dei concetti di *intersezionalità* e *assemblaggio*. Una preoccupazione simile, ma da prospettiva semiotica e articolata nel campo semio-linguistico, ci sembra essere riconosciuta da Zanelli, S. (2025). *Deleuze con Peirce. Un equivoco impossibile*. Orthotes, dove si discute l’*efficacia pragmatica* della formula deleuziana e guattariana di “*agencement* collettivo di enunciazioni”. Per noi è ancora una volta la performatività del concetto di *agencement* a fare problema e in ciò riconosciamo il suo valore teorico: se la relazione precede l’identità e la relazione è un atto che dura e muta e diviene, occorre abbandonare il campo ontologico (dominio dell’*essere*) per approdare a un campo pragmatico (dominio del *fare*).

Ignora ogni differenza tra l'artificiale e il naturale. Ignora la distinzione dei contenuti e delle espressioni, come quella delle forme e delle sostanze formate, che esistono solo mediante gli strati e rispetto agli strati" (Ivi, pp. 92-93).

Il piano di consistenza è il luogo di composizione degli operatori differenziali, dei piani vibranti e delle molteplicità, ciò che caratterizza la natura produttiva dell'assemblaggio, nella sua eterogeneità.

E più avanti: "Non c'è enunciato individuale, non ce n'è mai. Ogni enunciato è il prodotto di un concatenamento macchinico, cioè di agenti collettivi di enunciazione (per «agenti collettivi» non intendere popoli o società, ma le molteplicità). Ora, il nome proprio non designa un individuo: al contrario, è quando si apre alle molteplicità che lo attraversano da parte a parte, è all'uscita del più severo esercizio di spersonalizzazione che l'individuo acquista il suo vero nome proprio. Il nome proprio è l'appercezione istantanea di una molteplicità. Il nome proprio è il soggetto di un puro infinito compreso come tale in un campo di intensità" (Ivi, pp. 61-62).

Allora torniamo a Puar, la quale in un paragrafo dell'articolo sopra citato dal titolo "Cyborgs and Other Companionate Assemblages" utilizza, come abbiamo visto, il concetto deleuziano di *agencement* per proporre una revisione radicale del modo in cui pensiamo i corpi, le identità e le loro intersezioni, nella prospettiva di superare i limiti rappresentazionali e identitari che hanno caratterizzato gran parte della teoria femminista e postmoderna. Lo spostamento ontologico operato da Puar in merito al significato di *agencement* - non una semplice somma o giustapposizione di elementi, bensì un principio di organizzazione relazionale, una disposizione dinamica di forze, connessioni e intensità - le permette di allinearsi con quella costellazione di pensatrici che, a partire da Donna Haraway e Elizabeth Grosz fino a Karen Barad e Luciana Parisi, hanno messo in questione la centralità del linguaggio - meglio di una certa idea di linguaggio - e della rappresentazione nelle teorie femministe e post-strutturaliste. Tali autrici, che Puar definisce come "femministe materialiste", consapevole delle differenti posizioni interne, condividono il tentativo di riconoscere una vitalità della materia, ossia la capacità del materiale di agire, di produrre effetti, di partecipare ai processi di soggettivazione e di mondo.

Cita in proposito Karen Barad, che abbiamo già incontrato, "Language has been granted too much power. The linguistic turn, the semiotic turn, the interpretative turn, the cultural turn; it seems that at every turn lately every 'thing' is turned into language or some other form of cultural representation [...]. There is an important sense in which the only thing that does not seem to matter anymore is

matter” (Barad 2003, p. 801)²³⁰. La materia per Barad non è sostanza passiva su cui il discorso agisce, ma una forza performativa, un “fare” che partecipa alla costruzione del reale. L’*agencement*, in quanto composizione di forze e non semplice rappresentazione, diventa così il dispositivo teorico privilegiato per pensare la materia come relazione e il corpo come superficie di articolazione di elementi eterogenei – biologici, tecnologici, istituzionali, informativi. La stessa idea di *agencement* collettivi di istanze enuncianti proposta da Paolucci (2020), attraverso la mediazione di Deleuze e Latour, va in questa direzione.

Rispetto al tema della soggettività, di nostro interesse, l’*agencement* dissolve l’idea stessa di soggetto come punto fisso. Esso sostituisce alla logica del posizionamento quella della connessione, in cui le categorie identitarie non sono attributi o assi strutturali, ma eventi relazionali, emergenze temporanee all’interno di un campo di forze.

È all’interno di questa cornice che il concetto di cyborg acquista, in Puar, uno spessore teorico di facile scivolosità. E cita Dianne Currier²³¹ la quale ha messo come la figura del cyborg, pur concepita da Haraway per superare i dualismi moderni – umano/animale, organico/inorganico, naturale/artificiale, mente/corpo– finisca talvolta per reinscrivere il corpo in una logica binaria, trattando la tecnologia come un’“aggiunta” a un corpo umano preesistente. “In the construction of a cyborg, technologies are added to impact upon, and at some point intersect with a discrete, non-technological ‘body.’ [...] Thus, insofar as the hybrid cyborg is forged in the intermeshing of technology with a body, in a process of addition, it leaves largely intact those two categories—(human) body and technology—that preceded the conjunction” (Currier 2003, p.323), e proprio per questo il cyborg rischia di lasciare intatti i due poli che avrebbe dovuto destabilizzare. Rispetto a questa visione, l’*agencement* fornisce un correttivo fondamentale: non vi è un corpo e una tecnologia che si sommano, ma un processo di co-emergenza, in cui gli elementi biologici, tecnici, affettivi e istituzionali si costituiscono reciprocamente. Il cyborg non è, dunque, un soggetto ibrido nel senso di una fusione tra umano e macchina, bensì un campo di relazioni in cui si modulano capacità e potenze, un insieme instabile di connessioni che definiscono di volta in volta che cosa un corpo può fare, percepire, sopportare. E per tutto il resto c’è la delega di processi cognitivi e non. In conclusione del suo saggio dice Puar: “To return to the title of this piece, and the juxtaposition that Haraway (unfortunately, but presciently) renders, would I really rather be a cyborg than a goddess? The former hails the future in a teleological technological determinism—culture—that seems not only

²³⁰ Barad, Karen. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society* 28, no. 3: 801–831.

²³¹ Currier, Dianne (2003). Feminist technological futures: Deleuze and body/technology assemblages. *Feminist Theory* 3: 321–38.

overdetermined, but also exceptionalizes our current technologies. The latter—nature—is embedded in the racialized matriarchal mythos of feminist reclamation narratives. Certainly it sounds sexier, these days, to lay claim to being a cyborg than a goddess. But why disaggregate the two when there surely must be cyborgian goddesses in our midst? Now that is a becoming-intersectional assemblage that I could really appreciate” (Puar 2012, p. 63).

8.2 *Divenire (ancora più) cyborg*

What we make and what (we think) we are co-evolve together.

Katherine Hayles, *My Mother Was a Computer*.

Puar ci propone quindi una riformulazione del cyborg non più come una figura metaforica, né come un simbolo dell’epoca tecnologica, ma come una pratica di *agencement*: un dispositivo di articolazione e di misura che permette di leggere come le connessioni tra elementi eterogenei – infrastrutture digitali, corpi viventi, norme giuridiche, affetti e dispositivi – producano effetti materiali di vulnerabilità, di potenza o di esclusione. La preferenza per il cyborg segnava già il desiderio di sottrarsi ai dualismi moderni, ma la teoria dell’*agencement* consente di approfondirne le implicazioni, liberando il cyborg dal rischio di essere un nuovo soggetto, seppure ibrido, e restituendolo come metodo di composizione. Diventare cyborg, nell’accezione che Puar recupera da Deleuze e Guattari, non significa più incarnare un’identità tecno-umana, ma abitare un processo di montaggio, una zona di contatto in cui umano, non umano e tecnologico si articolano in configurazioni sempre provvisorie. L’*agencement* offre così al cyborg il suo habitat teorico più coerente: non un modello identitario, ma una pratica cartografica per pensare la materia come relazione, per descrivere e trasformare i dispositivi che producono corpi, capacità e mondi.

Quando proviamo a trasferire questa cassetta degli attrezzi teorica all’interno dell’analisi degli assemblaggi tra umano e macchine dotate di linguaggio cosa accade? Abbiamo davvero bisogno di ripensare radicalmente il significato di umano, di linguaggio, di pensiero, di cognizione. La critica del “femminismo neo-materialista”, citato da Puar, al linguaggio escludente la materialità come può ri-dirsi alla luce del fatto che la materia non umana in questo momento è dotata di linguaggio? Abbiamo bisogno di tutto il pensiero di un femminismo della tecnologia per farlo e nella nostra conoscenza, la prima ad aver compreso la portata epocale delle trasformazioni apportate dalle tecnologie a lei contemporanee è stata proprio Donna Haraway. Per questo riteniamo sia importante tornare alla sua idea di cyborg prima di liquidarla come una giustapposizione o sommatoria di entità differenti in natura: il corpo (umano) e le tecnologie (non umane). Ogni avanzamento in termini di

tecnoscienze sposta un passo più in là l'instaurazione di confini, linee di demarcazione nei modi della categorizzazione. Quindi oggi ci chiediamo che senso abbia l'opposizione naturale e artificiale alla luce delle macchine dotate di linguaggio?²³²

Riteniamo ci siano due modi principali in cui possiamo considerare inconsistente questa opposizione: il primo è quello illustrato da Kate Crawford in *"Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence"*, nel mettere a tema il problema dell'impiego di materie prime di difficile reperimento come base materiale del funzionamento dei sistemi di IA. Dice Crawford: "The tech sector's demand for rare earth minerals, oil, and coal is vast, but the true costs of this extraction is never borne by the industry itself. On the software side, building models for natural language processing and computer vision is enormously energy hungry, and the competition to produce faster and more efficient models has driven computationally greedy methods that expand AI's carbon footprint." (Crawford 2021, p. 15)²³³

Obiettivo di Crawford è quello di ricomporre i modi di intendere cosa sia l'Intelligenza Artificiale, mostrando come ciascuna angolazione attraverso la quale questa viene definita stabilisca una cornice entro la quale l'IA viene compresa, misurata, valutata e governata. Tra gli esempi riportati si trovano la pubblicità e il marketing, di interesse per i produttori dei beni di consumo per l'infrastruttura aziendale. Oppure quanti sono interessati alla capacità di decisione e di comportamento razionale dei modelli. Oppure ancora il caso di coloro che si interessano a specifiche tecniche algoritmiche e, di conseguenza, limitano il campo di indagine alla continuità/discontinuità a livello di progresso tecnico senza prendere in considerazione "il costo computazionale di tali approcci e il loro impatto profondo su un pianeta sotto pressione" (Ibidem).

Da qui la sua proposta di considerare l'IA come né artificiale né intelligente: "Piuttosto, l'intelligenza artificiale è sia incarnata che materiale, composta da risorse naturali, combustibili, lavoro umano, infrastrutture, logistica, storie e classificazioni. [...] In effetti, l'intelligenza artificiale per come la conosciamo dipende interamente da un insieme molto più ampio di strutture politiche e sociali. E a causa del capitale necessario per costruire l'IA su larga scala e dei modi per vederla

²³² Per un'analisi semiotica del concetto di natura rimandiamo a STANO, S. (2023) *Critique of Pure Nature*. Cham: Springer. ISBN (Hardcover): 978-3-031-45074-7, ISBN (Softcover): 978-3-031-45077-8, e-ISBN: 978-3-031-45075-4.

²³³ Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press; tr. It. Nè intelligente né artificiale. Il lato oscuro dell'IA. Bologna: Il Mulino (2021).

ottimizzata, i sistemi di IA sono in definitiva progettati per servire gli interessi dominanti” (Ivi, p. 16).

In questo tentativo di ricomposizione di lettura socio-tecnica che interroga l’intersezione tra tecnologia, capitale e governance Crawford vede l’unico modo per descrivere le implicazioni di “che cosa viene ottimizzato, per chi e chi è che decide” (Ivi, p. 17). Inoltre, la scelta di utilizzare la metafora dell’Atlante come prospettiva modulare e modulabile di analisi di un mondo, cessa a parere dell’autrice di essere metaforica nella misura in cui “il campo dell’IA è il tentativo esplicito di catturare il pianeta in una forma leggibile da un punto di vista computazionale” (Ivi, p. 18). In ciò si delinea una prensione colonialista, poiché l’IA rappresenta il desiderio di “essere l’atlante” del mondo, ossia un suo modo di dominazione.

Un secondo modo di affrontare la tensione semantica tra naturale e artificiale a proposito di macchine dotate di linguaggio e della loro capacità di produrre enunciati, presuppone di assimilare ciò che è biologicamente umano a ciò che è naturale e contrapporlo quindi manicheisticamente a un insieme indistinto di cose artificiali. Quest’approccio è stato esplorato a fondo da quella corrente del pensiero femminista che prende il nome di cyberfemminismo e che vede in Manifesto Cyborg²³⁴ di Donna Haraway una delle più compiute formulazioni: la soggettività cyborg è per Haraway “un organismo cibernetico, un ibrido di macchina e organismo, una creatura che appartiene tanto alla realtà sociale quanto alla finzione” (Ivi, p. 35). E aggiunge “Vorrei sostenere il cyborg come finzione cartografica della nostra realtà sociale e corporea, e come risorsa immaginativa ispiratrice di accoppiamenti assai fecondi” (Ibidem). Si tratta quindi di accoppiamenti che consentono di ripensare la riproduzione biologica e culturale fuori dall’eterosessismo e dal patriarcato, attraverso la costruzione di “realtà sociali e corporee in cui le persone non temano la loro parentela con macchine e animali insieme, né identità sempre parziali e punti di vista contraddittori”. (Ivi, p. 42)

Nella formulazione di Haraway la soggettività cyborg opera tre tipologie di sconfinamenti semioticamente definibili come paradigmatici:

- i) il primo quello tra umano e animale in cui e per cui, il linguaggio, l’utilizzo di strumenti, il comportamento sociale e alcuni eventi mentali non costituiscono più tratti nucleari di separazione.
- ii) il secondo sconfinamento è quello che regge la distinzione tra organismo (animale e umano) e macchina: la macchina vivente autopoietica di Varela e Maturana (1980) è

²³⁴ Haraway, D. J. (1991). *A Cyborg Manifesto*. Routledge; tr. it. *Manifesto Cyborg*. Donne, tecnologie e biopolitiche del corpo. Feltrinelli (2018).

esempio sintagmatico di ricombinazione delle unità concettuali componenti ciò che solitamente definiamo organismo. Nell'universo del discorso della Haraway questo si traduce nella "determinazione di protocolli tecnologici" nel senso della costituzione di uno "spazio ideologico aperto dalla riconcettualizzazione di macchina e organismo come testi codificati attraverso i quali partecipiamo al gioco di leggere e scrivere il mondo" (Ivi, p. 40).

- iii) infine la dissolvenza non-binaria del cyborg investe il confine tra il fisico e non-fisico che si manifesta nell'"ubiquità e nell'invisibilità" di queste macchine che si estende alla coscienza o alla sua simulazione, rendendo difficile un loro riconoscimento materiale e politico.

Il femminismo cyborg della Haraway si impegna a costruire un "noi" che non ricerca "nessuna matrice naturale di unità" e a ribadire che "nessuna costruzione è totalizzanti" (Ivi, p. 51), da qui una critica al marxismo e al femminismo socialista e radicale. Da un lato il marxismo prendeva in considerazione solo il lavoro produttivo salariato ("l'autocostruzione dell'uomo si basa sul dominio della natura" Ivi, p. 107), escludendo il lavoro di riproduzione come base della stessa riproduzione sociale antecedente al dominio della natura. A sua volta il femminismo socialista eredita la categoria del lavoro alienato come struttura ontologica dal pensiero marxiano, riportandolo per analogia sul lavoro di riproduzione. E ciò impedisce di dare il giusto peso a "L'unità delle donne si fonda qui su di un'epistemologia basata sulla struttura ontologica del "lavoro" (Ivi, p. 52). Per quanto riguarda la variante del femminismo radicale di MacKinnon, secondo Haraway essa rappresenta "una caricatura delle tendenze appropriatrici, incorporanti e totalizzanti delle teorie occidentali per le quali l'azione si fonda sull'identità" (Ibidem). In ballo c'è evidentemente l'idea di intersezionalità: "Non conosco nessun'altra epoca storica in cui ci sia stato un maggior bisogno di unità politica per affrontare seriamente le dominazioni di "razza", "genere", "sessualità" e "classe". Non conosco neppure un'altra epoca in cui il tipo di unità che noi potremmo contribuire a costruire sarebbe stato possibile. "Noi" non possediamo più la capacità simbolica o materiale per imporre a "loro" il nostro modello di realtà, o almeno "noi" non possiamo dirci innocenti della pratica di queste dominazioni. Le donne bianche, anche le femministe socialiste, hanno scoperto (o meglio, glielo hanno fatto notare tirandole per i capelli) la non-innocenza della categoria *donna*" (Ivi, pp. 50-51).

Ci sembra indicativo, allora, che Haraway definisca il cyborg una creatura di un mondo post-genere, riconfigurato dalle relazioni sociali di scienza e tecnologia. Come del resto dimostra oggi il superamento della visione per cui il genere sia solo un effetto discorsivo, convinzione di matrice costruttivista; tutto il neofemminismo materialista, che per comodità riconduciamo al pensiero di Rosi

Braidotti, va infatti nella direzione di rileggere il genere sempre come un assemblaggio material-discorsivo. Inoltre, “l’intero universo degli oggetti che possono essere scientificamente conosciuti deve essere formulato come un problema di ingegneria della comunicazione (Ivi, p. 58). L’informatica del dominio e dei sistemi controllati tramite feedback coinvolge non solo il mondo delle comunicazioni in senso stretto, ma anche la biologia; ed entrambe poggiano sull’elettronica e la microelettronica come “fondamento tecnico dei simulacri, cioè copie senza originali” o una teoria del testo (per coloro che oppongono resistenza). “Entrambe sono semiologie cyborg” (Ibidem).

Questo perché “il genere è un campo di differenza strutturato e strutturante dove i toni di una estrema localizzazione del corpo personalizzato e individualizzato intimamente, vibrano nello stesso campo delle emissioni globali ad alta tensione. La corporeità femminista non si riferisce a una località fissa in un corpo reificato, femminile o altro, ma a nodi nei campi, inflessioni negli orientamenti e a responsabilità per la differenza nei campi materiali e semiotici del significato. L’aver corpo è protesi significativa” (Ivi, p. 118).

L’ambivalenza dell’essere e avere corpo è stata attentamente messa a tema da Gallese e Morelli (2024)²³⁵ nel testo *Cosa significa esseri umani*. Nel discorso del neuroscienziato, ciò che consente la messa a fuoco del soggetto su di sé, cioè di fare esperienza di sé è l’opacizzazione dell’io operata dal linguaggio. Ancora con una metafora della visione Gallese descrive l’esperienza del soggetto, in una primigenia condizione mediata dal corpo, come non esserci, una condizione in cui l’io non è oggetto di attenzione, ognuno è totalmente trasparente, in quanto semplicemente angolo prospettico del mondo. Poi la metafora continua di notte, quando accendiamo la luce e guardiamo fuori dalla finestra, vediamo un albero in trasparenza, ma contemporaneamente ci vediamo mentre guardiamo l’albero. “Questa luce che accendendosi, dice Gallese, rende visibile l’oggetto è il linguaggio. È il linguaggio a rendere opaco il sé che guarda il mondo, trasformandolo nel primo pronome singolare, dando forma all’io” (Ivi, p. 58). Questo sé opacizzato attraverso il linguaggio, non è più integrato per mediazione del corpo nel mondo circostante, non è più quindi solo corpo, ma “ha un corpo”. Lo stesso linguaggio viene da Gallese assimilato a un “utensile cognitivo” che consente l’emergere dell’aspetto autoetico e dell’autoriflessione. “L’autoriflessione linguistica è affiancata e potenziata, moltiplicata da tutte le altre mediazioni tecnologiche inventate via via che l’evoluzione culturale progrediva” (Ivi, p. 60). Ad esempio, la rappresentazione del mondo nell’esternalizzazione dei disegni sulle pareti della grotta determina che non solo agiamo nel mondo, ma diventiamo spettatori della rappresentazione che abbiamo creato. Questo è un esempio di “opacizzazione tecnologicamente

²³⁵ Gallese, V., Morelli, U. (2024). Cosa significa essere umani? Corpo, cervello e relazione per vivere nel presente. Raffaello Cortina Editore

mediata". La mediazione tecnologica e simbolizzante non serve solo a fare qualcosa, ma diviene strumento per dire qualcosa a qualcuno e contemporaneamente dirsi. In questo senso ci sembra andare l'analisi di Cristina Voto (2021)²³⁶ quando si focalizza sulle trasformazioni del senso prodotte dai volti artificiali nelle *digital arts*. Dice qui Voto: "Opacizzare significa sottrarre, addensare la natura diafana della rappresentazione per inquadrarla, catturarla e offrirla alla vista", suggerendo così l'idea che uno sguardo che opacizza "addensa la trasparenza auratica della tecnologia digitale per restituire un'agentività all'intelligenza artificiale" (Ivi, p. 293). Intrevediamo qui in atto quasi un meccanismo di rovesciamento che permette alla macchina di soggettivarsi attraverso l'opacizzazione del linguaggio. Pensiamo, ad esempio, a quando gli algoritmi di machine learning vengono addestrati su dati sintetici generati da altri modelli generativi. La macchina intelligente dice se stessa e dice la rappresentazione del mondo da lei stessa mediata.

Anche in Kathrene Hayles, l'ibridazione umano-macchinico diviene strutturale. Con la sua idea di "assemblaggi cognitivi", già in *Unfinished Work. From Cyborg to Cognisphere*,²³⁷ Hayles (2006) provava a introdurre due tasselli per *finire il lavoro* iniziato da Haraway: innanzitutto introducendo il concetto di *Regime di Computazione*, ripreso da Thomas Whalen (2000)²³⁸ il quale definiva lo scambio di dati tra macchine e tra macchine e umani come un fenomeno globale chiamato "cognisphere". Queste le parole di Hayles: "Expanded to include not only the Internet but also networked and programmable systems that feed into it, including wired and wireless data flows across the electromagnetic spectrum, the cognisphere gives a name and shape to the *globally interconnected cognitive systems in which humans are increasingly embedded*. As the name implies, humans are not the only actors within this system; machine cognizers are crucial players as well. If our machines are 'lively' (as Haraway provocatively characterized them in the 'Manifesto'), they are also more intensely cognitive than ever before in human history" (Hayles 2006, p. 161). Nel 2006 le macchine dotate di linguaggio non avevano ancora fatto il loro ingresso in scena, ma già questa fisionomia cognitiva, anche in assenza di linguaggio, era ben evidente. Il secondo tassello a completamento del lavoro di Haraway è per Hayles rappresentato dal lavoro di Thompson e Varela che nel 2006 lei dichiarava essere in uscita con il titolo: *Why the Mind Isn't in the Head: The Lived Body in Biology*,

²³⁶ Voto, C. (2021). Opacizzare il volto artificiale attraverso le arti digitali: Errori, deformità, materia, intersoggettività". Lexia, Volti artificiali(37–38), 285–302. <https://doi.org/10.4399/978882553853314>

²³⁷ Hayles, N. K. (2006). *Unfinished Work: From Cyborg to Cognisphere*. *Theory, Culture & Society*, 23(7–8), 159–166. <https://doi.org/10.1177/0263276406069229>

²³⁸ Whalen, T. (2000). *Data Navigation, Architectures of Knowledge*. paper presented at the Banff Summit on Living Architectures: Designing for Immersion and Interaction, Banff New Media Institute, 23 Sept.

Cognitive Science and Human Experience per Harvard University Press.²³⁹ Al di là delle vicende editoriali, l'elemento di interesse per Hayles rispetto al loro lavoro era il seguente: "At the same time, advances in cognitive science, neurology and related fields of brain science are clarifying the neurological basis for human perception and subjectivity, leading to the possibility, [...] that the subject-object split institutionalized by the birth of modern science can at last be addressed by a cognitive science powerful enough to begin to explain the physical and psychological bases for human constructions of reality. These results validate Haraway's (1988) call for 'situated knowledges', demonstrating that there is no way to know the world except through the subjectivity that precedes and grounds our objective accounts" (Ivi, p. 162).

La Hayles abbraccia a suo modo l'idea che la mente non sia contenuta nella nostra scatola cranica e che i processi cognitivi che ci rendono così umani siano in realtà distribuiti in un assemblaggio con l'ambiente semiotico che ci circonda. La semiotica cognitiva questo lo sa bene (Paolucci 2021a; Lobaccaro 2022)²⁴⁰.

Citiamo infine il modo di intendere il concetto di cyborg proposto da Paolucci (2025c) nel suo lavoro proprio sul tema delle macchine dotate di linguaggio. Nati Cyborg vuole essere un omaggio al lavoro pionieristico di Andy Clark (2004)²⁴¹ Natural-born cyborg. Dice Paolucci: "la forma "cyborg": un essere che combina elementi organici e non-organici, meccanici o elettronici, con l'obiettivo di mostrare che questa forma cyborg – che innesta protesi e interfacce a cui deleghiamo parti fondamentali del nostro lavoro cognitivo – sia in forma profonda e non banale la natura stessa dell'essere umano. E lo sia fin dal principio e ben prima dell'intelligenza artificiale" (Ivi, p. 13-14). Perché? Perché il cyborg non è solo un prodotto di un certo grado di sviluppo delle tecnologie (digitali e wireless, aggiungiamo), ma inerisce nel profondo la natura umana? Paolucci risponde a questa

²³⁹ Tuttavia, nelle bibliografie a nostra disposizione non è stato reperibile; abbiamo però trovato un contributo²³⁹ sempre di Varela e Thompson in cui loro stessi citano questo libro in Forthcoming Varela, F. J., & Thompson, E. (2000). Neural Synchrony and the Unity of Mind: A Neurophenomenological Perspective. Citiamo dalla copertina: "FV passed away May 28, 2001. This paper is based on an invited lecture he gave at the ASSC Meeting at

Brussels in June 2000, and on subsequent extensive revisions in collaboration with ET during the period of December 2000 to March 2002. It was then presented by ET at the conference "Beyond the Hard Problem: Consequences of Neurophenomenology," at Boulder, Colorado, in May 2001. ET is supported by the Social Sciences and Humanities Research Council of Canada, the McDonnell Program in Philosophy and the Neurosciences, and the Fetzer Institute through a grant to the Center for Consciousness Studies, University of Arizona, Tucson".

²⁴⁰ Lobaccaro, L. (2022) La semiotica cognitiva in Italia. Genesi, diramazioni, evoluzioni., in: Cura del senso e critica sociale. Ricognizione della semiotica italiana, Roma, Mimesis Edizioni srl, pp. 61 - 94

²⁴¹ Clark, A. (2003). Natural-born cyborgs: Minds, technologies, and the future of human intelligence. Oxford University Press.

domanda raccontando qualcosa che può non provocare un certo disturbo, poiché erode nel profondo il senso di superiorità che caratterizza una certa idea di umano: “l’intelligenza artificiale, nonostante i suoi risultati straordinari, è debole e rimanda la sua stessa essenza al concatenamento con l’essere umano e al cyborg che crea con esso”. Per Paolucci tanto l’IA quanto l’umano sono “imbecilli” per natura: “la parola “imbecille” deriva infatti dal latino in-bacillum, che è il diminutivo di “baculum”, cioè “senza bastone”. Con questo termine si indicava originariamente una persona debole, fragile, debole proprio perché “senza bastone”, pur avendone evidentemente bisogno” (Ivi, p. 69). E altrove Paolucci cita Clark a proposito del concetto di mente estesa e delle resistenze più radicali a concepire la debolezza costitutiva della specie umana che è stata condizione di un certo modello di evoluzione di specie – ricordiamo infatti che gli animali umani sono una delle specie neoteniche per eccellenza, in cui lo sviluppo del cervello del cucciolo di uomo avviene per i due terzi al di fuori dell’utero materno e dunque intensamente accoppiato con l’ambiente esterno e, al contempo, completamente vulnerabile. Ecco in cosa consiste questa delega all’ambiente per Clark: “Da anni sostengo che noi esseri umani siamo – e siamo sempre stati – “menti estese”: sistemi di pensiero ibridi, definiti (e continuamente ridefiniti) da un ricchissimo mosaico di risorse, solo una parte delle quali risiede nel cervello. Sembriamo sorprendentemente restii a riconoscere la nostra natura ibrida. Da tempo coltiviamo un’immagine molto restrittiva di noi stessi, un’immagine che alimenta il timore che nuovi strumenti e tecnologie producano un impoverimento mentale. Nel Fedro di Platone, datato intorno al 370 a.C., troviamo un’esposizione molto chiara di questa paura: l’idea che invenzioni allora “avveniristiche” come lettura e scrittura avrebbero avuto effetti catastrofici sulla memoria. Si temeva che quelle innovazioni generassero menti pigre, facendoci credere di conoscere più di quanto in realtà sapessimo grazie a mezzi di archiviazione esterni; economici, ma superficiali. Oggi quelle ansie suonano quasi ridicole. Ma non abbiamo ancora reciso la loro radice: la nostra idea erronea dei fenomeni cognitivi. È singolare, specie se si considera che la storia umana è scandita, più volte, dagli effetti trasformativi di nuovi strumenti e pratiche. Si potrebbe anzi sostenere che sia un tratto fondamentale della nostra specie distribuire in questo modo il carico cognitivo, diventando ciò che ho definito “natural-born cyborgs”.²⁴²

La delega al mondo esterno dei processi cognitivi non è effetto di pigrizia, quindi, ma evoluzione di specie. Proviamo quindi a vedere cosa accade ad alcuni predicati che abbiamo definito soggettali se assumiamo l’ibrido umano-macchinico.

8.3 Predicati soggettali: pensare, parlare, produrre significato

²⁴² Clark, A. (2025). Extending Minds with Generative AI. *Nature Communications*, 16(1), 4627. <https://doi.org/10.1038/s41467-025-59906-9>;

Se, come affermato nel paragrafo precedente, la nostra prima intenzione era mettere in discussione l'opposizione tra naturale e artificiale nella direzione di un tentativo seminale di definizione dell'enunciazione macchinica, occorre ora ripartire dall'interno di queste due categorie per evidenziare quali nuovi *assemblaggi significanti* ci consentano di proseguire la riflessione.

8.3.1 Una variazione sul tema del pensare

Il celebre saggio di Turing²⁴³ (1950), *Computing Machinery and Intelligence* fornisce un primo slittamento semantico del lessema intelligenza, quando riconfigura il tentativo di risposta alla domanda “le macchine possono pensare?” nei termini di un gioco di imitazione. Un gioco che — è difficile non notarlo — propone originariamente una definizione del maschile e del femminile come effetti di senso, e non come essenze predeterminate.

Riprendiamo da Paolucci (2025c) la descrizione del test di Turing: “Nel gioco dell'imitazione, un interlocutore umano deve infatti indovinare chi è la femmina, che cerca di aiutarlo, e chi è il maschio, che cerca invece di ingannarlo. Se prima e dopo la sostituzione dell'uomo con una macchina, la percentuale delle volte in cui l'osservatore umano indovina chi cerca di aiutarlo e chi cerca di ingannarlo rimane simile, allora, secondo Turing, la macchina dovrebbe essere considerata intelligente e dotata di pensiero” (Ivi, pp. 101-102).

È nel significato supposto del femminile, veicolato attraverso un gioco di simulazione ingannevole, che il giocatore maschio tenta di nascondere la propria identità maschile e così depistare l'interrogatore. Il ruolo della giocatrice, invece, è quello di convincere in ogni modo l'interrogatore della veridicità della propria identità. Su questo modello dell'inganno, che è un meccanismo profondamente semiotico, si fonda anche l'attribuzione di intelligenza nel caso in cui al posto dell'uomo si trovi la macchina, la cui intelligenza è oggetto di prova.

Alcune rapide considerazioni a partire da questo scenario:

- i) Come osservano Drage e Frabetti (2023)²⁴⁴, l'articolo di Turing stabilisce una curiosa coincidenza tra l'indovinare correttamente il genere (la distinzione uomo/donna) e l'indovinare correttamente l'umanità intelligente (la distinzione umano/artificiale). Tale

²⁴³ Turing, A. M. (1950). *Computing Machinery and Intelligence*. *Mind, New Series*, 59(236), 433–460.

²⁴⁴ Drage, E., & Frabetti, F. (2023). *AI that Matters: A Feminist Approach to the Study of Intelligent Machines*. In J. Browne, S. Cave, E. Drage, & K. McInerney (A c. Di), *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines* (p. 0). Oxford University Press. <https://doi.org/10.1093/oso/9780192889898.003.0016>

coincidenza, in realtà, non è così curiosa se si considera l'omosessualità di Turing e la sua storia personale.²⁴⁵ Una delle principali teoriche queer, Judith Butler, sottolinea come l'accesso al sociale — l'ingresso dell'umano nell'intelligibile — sia mediato dal genere assegnato alla nascita, e dunque che il gioco dell'imitazione, anche quando coinvolge una macchina, è superato se il genere viene correttamente attribuito in termini performativi.

- ii) Una seconda riflessione deriva invece dall'ontologia dipendente dal valore della performance cognitiva descritta da Turing. Come evidenzia Blair Attard-Frost (2021)²⁴⁶, “Favouring human intelligence as the ideal, default model for AI in all situations encourages the designers and developers of AI systems to evade critical analysis of the values and biases that often underwrite human cognition and human decision-making” (Ivi, p. 34) E aggiunge: “Turing believes it could be technically possible to design a machine for the simple purpose of enjoying the taste of a dessert, but he dismisses any attempt to make such a machine as “idiotic” (1950, 448). Turing dismisses the idea of the dessert-eating machine not because of any technical impossibility, but because of a perceived lack of utility value in the cognitive activities associated with dessert-enjoying” (Ibidem). A partire da questa osservazione, Attard-Frost invita a reimmaginare l'intelligenza e di conseguenza l'IA, decentrando l'umano dalla relazione IA-umano e portando una serie di esempi che varrebbe la pena esplorare per ridefinire il rapporto essere umano-macchina intelligente. Ci riferiamo ad esempio a progetti in corso su buone prassi de/post-coloniali di addestramento dei sistemi di IA. La stessa riflessione sui bias di cui abbiamo parlato in 3.2 attiene a questo passaggio; e tuttavia ribadiamo che se non cambia concettualmente il modo di intendere la relazione umano-macchina intelligente, non basteranno le mitigazioni dei bias a mettere in discussione un modello relazionale ed epistemico basato sull'estrattivismo e sullo sfruttamento.²⁴⁷

8.3.2 Due variazioni sul tema del parlare

²⁴⁵ Su questo si veda il bel fumetto di Paggiaro, A. e Riccioni F. (2012). *Enigma. La strana vita di Alan Turing*. Rizzoli-Lizard. Ringrazio ancora il collega Marco D'Alessandro per questo suggerimento.

²⁴⁶ Attard-Frost, B. (2024). *Queering intelligence. A theory of intelligence as performance and a critique of individual and artificial intelligence*. In *Queer Reflections on AI. Uncertain Intelligences*. Routledge.

²⁴⁷ Si veda in proposito Goodman, A. & Ippolita (2025). *Macchine neurodivergenti. Relazioni postumane e algoritmi queer*. Bologna: Ombre Corte.

I modelli generativi di IA sono in grado di creare testi; la loro infrastruttura tecnica permette di simulare una conversazione tra esseri umani. La reazione a questa capacità tecnica ha definito gli schieramenti in campo, con diverse sfumature, tra:

- i) chi ritiene che si tratti di una pura manipolazione del significante, in completa assenza di significato (una probabilità condizionata associata alla potenza di calcolo);
- ii) e chi sostiene invece che il processo di generazione testuale di un LLM non differisca in modo sostanziale da ciò che accade nel nostro cervello quando parliamo o scriviamo.

Ancora una volta, è in gioco il valore semantico, in questo caso, del verbo parlare. Cosa significa “parlare”?

Negli anni Ottanta, Gayatri Spivak,²⁴⁸ in aperta critica con la concezione deleuziana della soggettività desiderante e del soggetto biopolitico di Foucault, affermava l’incapacità del soggetto coloniale — il grande Altro dell’Europa — di parlare per sé.

Come osserva Cristina Demaria, “il colonialismo fu anche un’“operazione del discorso” che interpellava i soggetti coloniali, inscrivendoli in un sistema di rappresentazione, sebbene essi ne fossero già parte” (Demaria 2019). Per Spivak, “parlare” implica due significati di rappresentazione: da un lato, *representation* nel senso politico di “parlare per”, ciò che avviene nella delega e nelle forme della rappresentanza decisionale. Dall’altro ogni rappresentazione implica una *re-representation* nel senso estetico o filosofico di “ripresentare”, ossia un presentare di nuovo in forma mediata e, dunque, anche risignificata. In ogni passaggio di mediazione avviene qualcosa che semioticamente ridefinisce il perimetro del contenuto che viene mediato. Come negli esempi di traduzione di cui abbiamo parlato nel capitolo tre del presente elaborato, qualcosa sempre si perde e/o qualcosa sempre si aggiunge. Al soggetto coloniale entrambe queste forme sono precluse: la rivolta (*insurgency*) prende il posto dell’enunciazione (*utterance*), poiché non può esserci una presa di parola *autentica* e uno spazio di discorso è precluso. Il soggetto coloniale non parla, se non essendo parlato dalla lingua del soggetto colonizzatore al cui mondo deve adattarsi e al cui interno deve scomparire. L’analisi di Fanon nel primo capitolo di *Pelle nera, maschere bianche* racconta proprio in questi termini il rapporto tra “il nero e il linguaggio”.

Continuando a dare una cornice di riferimento per comprendere la non “neutralità” e “oggettività” del valore semantico del lessema |parlare|, ci spostiamo all’interno di una prospettiva postumanista e

²⁴⁸ Spivak, G. C. (1988). *Can the Subaltern Speak? Marxism and the Interpretation of Culture*, 271–313.

di linguistica applicata. Pennycook (2018)²⁴⁹ propone di ripensare il rapporto tra linguaggio, rappresentazione e materialità, poiché “there is no longer a world ‘out there’ separate from humans and represented in language but rather a dynamic interrelationship between different materialities” (Ivi, p. 449). Il linguaggio e la cognizione non sono proprietà inerenti o interiorizzate negli individui umani, ma pratiche semiotiche distribuite, multimodali e multisensoriali del quotidiano, che comprendono le relazioni dinamiche tra risorse semiotiche, attività, artefatti e spazio. Ancora una volta assemblaggi sugli strati. Infatti, questo approccio richiama immediatamente l’idea di assemblaggio di istanze enuncianti propria di una teoria dell’enunciazione impersonale, traspunta dal livello dell’apparato formale dell’enunciazione a quello dell’enunciazione enunciata. Inoltre, richiama l’idea di Eco secondo cui i segni non sono altro che espressioni materiali (cfr. *Trattato di semiotica generale*, p. 325).

8.3.3 Una variazione sul tema della produzione di significato

Sempre nell’ambito del post-umanesimo, ma con enfasi sulla continuità tra cognizione biologica e computazionale, Katherine Hayles (2019)²⁵⁰ introduce il concetto di *intermediazione*, con cui si riferisce all’interazione tra livelli dinamici di organizzazione, chiamati *media*; questi sono pattern quasi stabili che vengono catturati a un livello superiore per formare entità nuove e più complesse.

Hayles scrive: “The change of media in intermediation is crucial because it is what “locks in” lower-level dynamics and enables emergence to continue through successive levels of increasing complexity.” (Ivi, p. 33).

Questa struttura stratificata, i cui strati interni comunicano attraverso i segni, garantisce uno scambio bidirezionale, top-down e bottom-up: l’intermediazione descrive un’organizzazione ricorsiva che Hayles definisce eterarchia dinamica (*dynamic heterarchy*). L’intermediazione assume che i segni possano operare secondo dinamiche simili sia nella cognizione biologica sia nei media computazionali (dagli operatori logici ai livelli di codice sovrapposti). Riferendosi a Deacon e Wheeler, Hayles spiega che i quadri teorici che essi sviluppano per l’interpretazione, lo scopo e il significato nell’ambito biologico possono, con le dovute modifiche, costituire un potente insieme di strumenti per comprendere come i computer producano significato e come le interazioni uomo-macchina abbiano trasformato radicalmente la nostra capacità di agire sul mondo, nel bene e nel male.

²⁴⁹ Pennycook, A., 2018. *Posthumanist applied linguistics*. Routledge.

²⁵⁰ Hayles, N. K. (2019). Can Computers Create Meanings? A Cyber/Bio/Semiotic Perspective. *Critical Inquiry*, 46(1), 32–55. <https://doi.org/10.1086/705303>.

Secondo Hayles, lo stesso meccanismo di interpretazione — ispirato a Peirce e inteso come “an action or behavior that has been consistently linked to an interior or external sign relation established through a historical connection” (Ivi, p. 44) — opera anche all’interno del computer, con lo stesso potenziale d’errore dovuto a fonti d’incertezza (errori di segnale, decadimenti, raggi cosmici, o l’imprevedibilità degli input nei programmi eseguibili). Alla tipica obiezione di tipo searliano, relativa all’assenza di *aboutness* e all’incapacità di riconoscere i fenomeni di assenza che governano la biologia, Hayles risponde affermando che anche i computer possiedono un orizzonte di mondo (*Umwelt*), di cui portano una certa quantità di informazione all’interno dei loro sistemi integrati.

Hayles riconosce l’urgenza di identificare prospettive alternative che riconoscano il contributo di altre specie alla semiosfera planetaria, così come quadri teorici capaci di sottolineare l’importanza dei media cognitivi nella creazione dei significati che guidano l’azione, la percezione e le decisioni ibride umano-tecniche nel mondo contemporaneo della cognisfera (Hayles, 2019, p. 52). L’idea del “adjacent possible” elaborata da Kauffman in relazione all’evoluzione biologica viene così ampliata: “When there is a very large space of possibility, he argues, the relevant issue is not the space’s theoretical size but the strong likelihood that pathways adjacent to the existing one will be followed, giving a teleological impetus to evolutionary developments.” (Ivi, p. 54).

Alla luce di questi slittamenti semantici, riteniamo che buona parte dei predicati che hanno costituito una certa idea di soggettività, siano stati già e possano essere ancora questionati come proprietà esclusive di un umano che alla prova dei fatti si frammenta in una molteplicità di forme di vita dotate di vari gradi di libertà, volontà e trasparenza rispetto alle facoltà di pensare, parlare e creare.

9.0 Filiera formativa STEM e questioni di genere ad essa connesse

Non sono quel tipo di persona che sa cosa fare
della propria vita e non so se lo sarò mai.
So però quello che mi piace, quello che potrebbe piacermi
e so per certo quello che non mi piacerà mai.

Anonima

9.1 Il genere come categoria analitica dello spazio-evento

Vorremmo aprire quest'ultimo capitolo con una dichiarazione di principio: parlare di questioni di genere - o come li definisce Judith Butler (1999) *gender troubles*,²⁵¹ con maggior efficacia espressiva - apre un campo di tensione epistemologica e politica fortemente perturbato. Intendiamo che i *problemi* teorici che crea la categoria di genere, difficilmente isolabile da altre marche identitarie come vedremo, ci espone alla necessità di una serie di operazioni: fornire definizioni, mettere in discussione assunti e narrazioni, risalire alle radici della nostra soggettività, del nostro essere soggetti sociali riconosciuti, ma soprattutto riconoscibili. Abbiamo, quindi bisogno di dare una definizione di *gender trouble* e abbiamo allo stesso modo bisogno di dire cosa la semiotica, come campo di studi che indaga il senso attraverso le sue mediazioni e traduzioni – non ci stanchiamo di ripeterlo – possa dire su questi *trouble*.

Partiamo dunque con le definizioni perché proviamo a mitigare gli equivoci propri del comunicare e del significare. Butler mobilita l'ambivalenza valoriale del termine "trouble", che nella versione italiana del suo libro è tradotto con "questioni" e non con problemi. Perché? Ci dà una giustificazione la traduttrice, Sergia Adamo, in nota alla Prefazione del 1999: "Tanto del significato e della vita di questo testo è racchiuso nella felice scelta del titolo inglese: Gender Trouble. Un «trouble» che non è un neutrale problema, o semplicemente una questione aperta, ma che mette in gioco il senso di un disturbo, di un fastidio, di un «guaio» e di una confusione, che si può provocare o subire, ma anche provocare e subire allo stesso tempo. La «questione di genere» disturba, dà fastidio, irrita, è urticante, ma allo stesso tempo può sparigliare le carte del suo stesso gioco. Non è insomma una frivola «questione» tra le tante. La plurivocità semantica e non conciliante che il titolo proietta sulle sue possibili letture è importante che resti come chiave d'accesso, al di là delle scelte traduttive e al di là della vita singolare che questo testo ha continuato ad avere in contesti a volte non previsti e mai del tutto prevedibili. Per questo va ricordato che questa traduzione italiana è in realtà una ri-traduzione,

²⁵¹ Butler, J. (1999). *Gender trouble: Feminism and the subversion of identity* (10. anniversary ed). Routledge; tr. it. Butler, J. (2017). *Questione di genere: Il femminismo e la sovversione dell'identità* (S. Adamo, Trad.). GLF editori Laterza.

che prova però a confrontarsi solo ed esclusivamente con il testo inglese e che nell'accettare la sfida si augura che ciò che ne sarà in italiano di Gender Trouble continui a provocare disturbi, problemi e confusione creativa (N.d.T.)" (Ivi, p. V).

Sarebbe in qualche modo a dire che laddove si parla di genere si parla di un elemento problematico o disturbante; foricamente pericoloso o rassicurante, polemico o pacifico (almeno per coloro che accettano una rappresentazione binaria del sistema sesso/genere), in ogni caso polarizzante. Genere è categoria analitica, così definita nel vocabolario online Treccani (di cui riportiamo solo alcune voci)²⁵²:

“gènere s. m. [dal lat. *genus* -nēris, affine a *gignēre* «generare» e alle voci gr. γένος «genere, stirpe», γένεσις «origine», γίγνομαι «nascere»]. – 1. Nel suo sign. più ampio, termine indicante una nozione che comprende in sé più specie o rappresenta ciò che è comune a più specie. [...] 2. a. Nel linguaggio com., l'insieme dei caratteri essenziali per cui una cosa è simile ad altre o differisce da altre; con tale sign., è per lo più sinon. di qualità, sorta, tipo, e anche di specie. [...] b. Con sign. più ampio, il g. umano, l'umanità intera, gli uomini nel loro complesso; in questo senso, anche la specie umana (che ha però sign. più restrittivo, e fa più sentire la contrapposizione agli animali). [...] 4. a. Categoria grammaticale esistente nelle lingue indoeuropee, semitiche e in molte altre famiglie linguistiche, alcune delle quali distinguono tre generi, maschile, femminile e neutro (per es., il latino, il greco, il tedesco), altre, come l'italiano e il francese tra le lingue moderne, soltanto due, maschile e femminile; la distinzione del genere, che solo in un ristretto gruppo di sostantivi è connesso con il genere naturale, si manifesta nella declinazione dei sostantivi, dei pronomi e degli aggettivi, e nell'accordo tra essi. b. Per estens., con riferimento alla specie umana, carattere maschile o femminile dell'individuo, anche in senso biografico, sociale, professionale, come nell'espressione identità di genere, con cui s'intende la costellazione di caratteri anatomo-funzionali, psichici, comportamentali che definiscono il genere in sé stesso e in quanto posseduto, accettato e vissuto dall'individuo nella storia familiare da cui proviene e nella società in cui vive. c. G. del verbo, categoria grammaticale della coniugazione, di natura sintattica; nelle lingue indoeuropee non ha rilievo morfologico, ma consiste solo nella possibilità che ha ogni verbo di accordarsi sintatticamente con i nomi: così il genere del verbo, e il verbo stesso, è transitivo, quando ha la possibilità di accordarsi con un nome in accusativo (o complemento diretto), intransitivo quando il caso o complemento è diverso dall'accusativo, oppure il verbo è usato assolutamente”.

²⁵² Per la definizione e le accezioni complete si veda: *Gènere—Significato ed etimologia—* Vocabolario Treccani. Recuperato 29 ottobre 2025, da <https://www.treccani.it/vocabolario/genere/>

Il campo semantico del termine genere di per sé fa perdere punti di riferimento e confini, poiché sembra avere come obiettivo quello di astrarre dalle differenze e dalle molteplicità per identificare e circoscrivere in categorie.

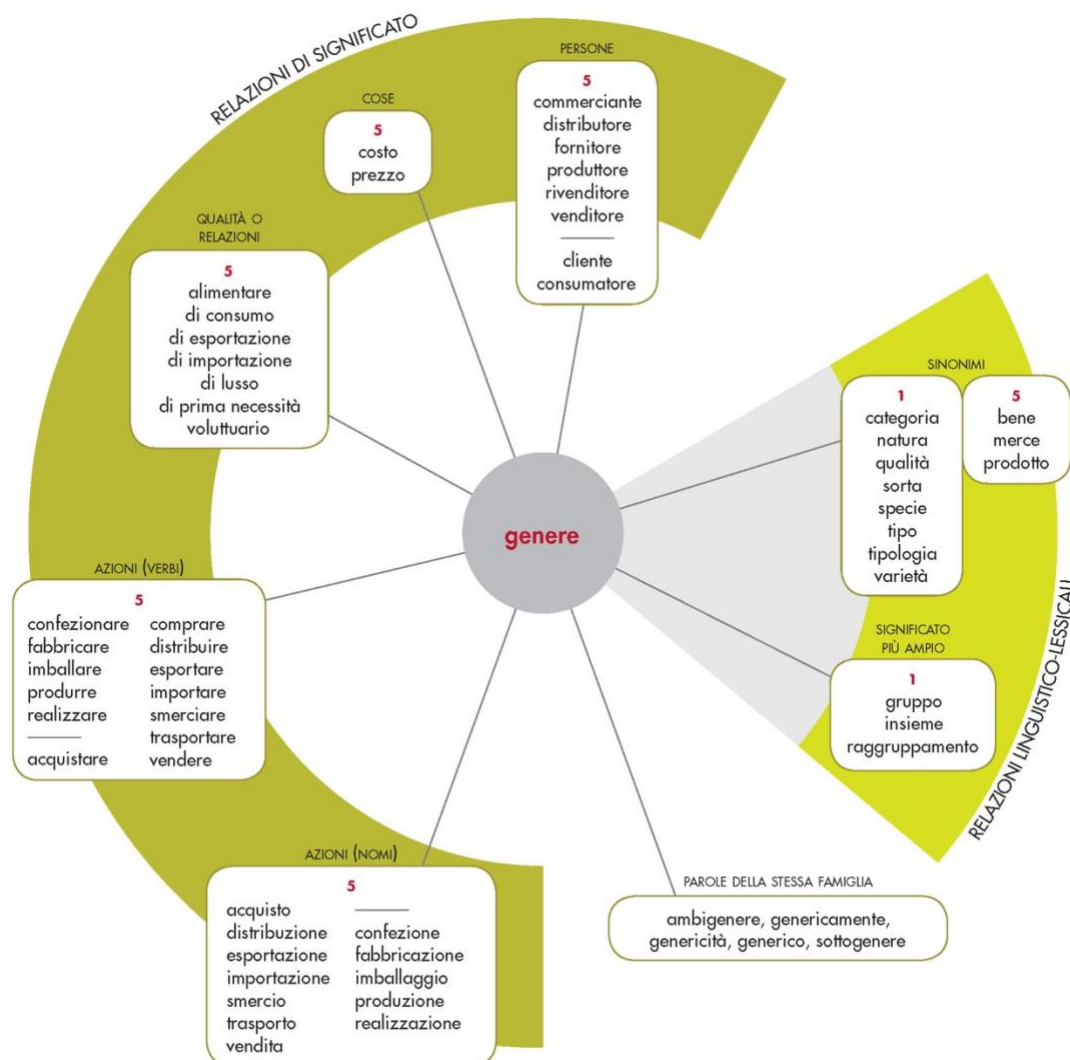


Figura 27: Mappa delle relazioni linguistico lessicali del termine “genere”²⁵³

La radice etimologica richiama alla potenza generatrice del genere, e ciò che viene generato è paradossalmente il genere stesso, nell’accezione per estensione con riferimento alla specie umana. È concetto di difficile definizione prima ancora che si estenda ai *gender trouble* su cui Butler si esprime e che troviamo a fondamento della nostra cultura e tradizione di pensiero già dai tempi della filosofia greca.

²⁵³ *Thesaurus Traccani* 2018

Partiamo da qui allora. Già in Platone, *Filebo* 12e sgg., si affermava; “Dal punto di vista del genere, ogni figura è una cosa sola, ma fra le parti del genere, rispetto alle parti, alcune sono al massimo grado opposto tra loro, alcune accade che siano diverse, e ciò può essere secondo un numero enorme di modalità”. Ma è con Aristotele che il genere trova la sua definizione logica: “Genere è ciò che si predica secondo l’essenza di molti che differiscono specificatamente” (*Topici*, 102a, 31 sgg.).

Il genere presiede a un gesto di astrazione delle differenze e del molteplice, attraverso qualcosa che potremmo chiamare una discretizzazione del continuo e che è fondamento di un certo modo di concepire la realtà delle cose in categorie o gruppi (per somiglianza e per differenza) propria certamente di una cultura occidentale.

Da una prospettiva semiotica, sviluppare analisi che tengano conto del genere significa cogliere criteri di classificazione per dare origine a tassonomie. Leggiamo infatti in Greimas e Courtès: “Lo studio della teoria dei generi, caratteristica di una cultura* (o di un’area culturale) data, è interessante solo nella misura in cui può mettere in evidenza l’assiologia* soggiacente alla classificazione; va inoltre comparata alla descrizione di altre etno- o socio-tassonomie” (Greimas e Courtés 1979, p. 144).²⁵⁴ L’opera di classificazione per genere e specie lascia emergere tutto il portato ideologico di un gesto fondativo del pensiero; questo è di indubbio interesse semiotico, anche se la “teoria dei generi” (dove genere è una classe di discorsi identificabile per la loro natura sociolettale e letteraria) per i due autori è di dubbio valore scientifico poiché poggia su un relativismo culturale. Il genere che presiede alla definizione specie-specifica degli elementi costituenti una molteplicità è semioticamente interessante poiché, in seno a una data cultura, è effetto dell’insieme di valori che sottendono un certo sistema di classificazione. Il genere, così inteso, pone costantemente un vincolo produttivo alla fondazione della soggettività in semiotica.

Quando circoscriviamo il concetto di genere al suo valore di categoria rappresentativa della differenza sessuale e dunque di un aspetto costitutivo della soggettività, il quadro sembra complicarsi. Infatti, spiega Demaria (2003)²⁵⁵:

“Indagare la costruzione delle soggettività contemporanee e gli effetti delle loro rappresentazioni vuol dire inoltre già applicare un’ottica semiotica, dal momento che i processi attraverso cui il soggetto si costituisce in quanto sessuato sono innanzitutto testuali e mediali (oltre che ovviamente

²⁵⁴ Greimas, A. J., Courtés, J., (1979). *Sémiotique. Dictionnaire raisonné de la théorie du langage*; tr. It. (2007) *Semiotica: Dizionario ragionato della teoria del linguaggio*. B. Mondadori. https://books.google.it/books?id=gXd-FU_xNI8C

²⁵⁵ Demaria, C., Tiralongo, A. (2003). *Teorie di genere. Femminismi e semiotica*. Milano: Bompiani.

sociali e culturali: ma l'uno non esclude l'altro). "Testualità" è quindi da intendersi in senso ampio, come insieme di valori e processi di significazione, meccanismi di senso che definiscono anche le pratiche, i rituali, gli spazi urbani, e i testi mediali e transmediali" (Ivi, p. 13). Demaria ci sembra dire che l'impresa semiotica ha per vocazione lo studio di quegli oggetti che fanno la nostra soggettività e tali oggetti semiotici sono gli enunciati, così come gli abiti – peirceanamente intesi - che incorporano strutture di potere denunciate dal femminismo. Se, come abbiamo più volte dichiarato, abbracciamo una teoria dell'enunciazione impersonale ed evenemenziale (Paolucci 2020), non possiamo che assumere come il genere insieme alla soggettività sia posto nell'atto enunciativo nello stesso momento in cui gli enunciati, che popolano la vita culturale dell'Enciclopedia, sono anch'essi prodotti. Nell'atto di enunciazione sono posti contemporaneamente il soggetto e l'enunciato. Possiamo concepire dunque il genere come quell'*effetto a priori* che per Paolucci descrive il che cosa dell'Enciclopedia. Ancora, ci sembra di poter dire che la categoria di genere è per eccellenza concetto enciclopedico e per nulla dizionariale. Citiamo ancora Demaria (2003):

“Il genere è dunque innanzitutto una categoria impiegata nell'indagine dei ‘caratteri’ che definiscono il modo in cui l'appartenenza a un sesso non solo è vissuta, ma anche trasmessa da istituzioni sociali (quali la famiglia, la scuola, i vecchi e nuovi media), per sottolineare come i significati, le pratiche e i modelli di comportamento legati all'assunzione di un'identità femminile o maschile - o oggi LGBTQ+ - non dipendano da caratteristiche immutabili e naturali, bensì dalla storia e dalle trasformazioni sociali e culturali” (Ivi, p. 17).

Ciò, in parte, si collega a quanto affermavamo in 1.1 sulle *situated knowledge* di Haraway²⁵⁶; tuttavia qui facciamo un passo avanti con Puar,²⁵⁷ nella misura in cui situarsi, posizionarsi, specificare il *da dove* parla *chi* parla è già l'effetto di una serie di mediazioni, spostamenti, eventi. In questo suo essere vissuto e trasmesso il genere interseca altri “differenziali di potere” (Ivi, p. 21) che agiscono nel virtuale enciclopedico e che sono ben rappresentati da quegli incorporali (schemi, norme, modelli, abiti, apprezzamenti collettivi, istituzioni) che si incorporano a seguito di particolari configurazioni compositazionali: ci riferiamo alla classe, alla razza, all'orientamento sessuale, al sesso e alla sessualità, all'età, alla dis-abilità.²⁵⁸ Il modo di composizione di questi “differenziali di potere” è problema che Puar legge attraverso il rapporto tra intersezionalità e assemblaggio: “One of Crenshaw’s foundational examples—that of the traffic intersection does indeed describe intersectionality as an

²⁵⁶ Haraway, D. (1988), Op. Cit.

²⁵⁷ Puar, J. (2012), Op. cit.

²⁵⁸ Demaria, C. (2016). Intersezionalità e femminismo transnazionale tra costruttivismo, post-strutturalismo e ‘performance’ epistemologiche. *Scienza & Politica. Per una storia delle dottrine*, Vol 28, No 54 (2016). <https://doi.org/10.6092/ISSN.1825-9618/6219>

event. [...] As Crenshaw indicates in this description, identification is a process; identity is an encounter, an event, an accident, in fact. Identities are multicausal, multidirectional, liminal; traces aren't always self-evident. The problem of how the two preexisting roads come into being notwithstanding, there is emphasis on motion rather than gridlock, on how the halting of motion produces the demand to locate.” (Puar 2012, p. 59). Eppure, questa dinamica dell'evento descritta da Crenshaw²⁵⁹ resta per Puar inscritta nella logica dell'identità. Alla logica dell'identità Puar contrappone l'assemblaggio dello *event-space* che riprende dal lavoro di Massumi, *Parables for the virtual: Affect, movement, sensation*.²⁶⁰ In particolare dalla descrizione di questo spazio-evento che interessa un presunto incremento di violenza nello spazio domestico durante il Super Bowl Sunday:

“The home entry of the game, at its crest of intensity, upsets the fragile equilibrium of the household. The patterns of relations between household bodies is reproblematicized. The game event momentarily interrupts the pattern of extrinsic relations generally obtaining between domestic types, as typed by gender. A struggle ensues: a gender struggle over clashing codes of sociality, rights to access to portions of the home and its contents, and rituals of servitude. The sociohistorical home place converts into an event space. The television suddenly stands out from the background of the furnishings, imposing itself as a catalytic part-subject, arraying domestic bodies around itself according to the differential potentials generally attaching to their gender type. For a moment, everything is up in the air—and around the TV set, and between the living room and the kitchen. In proximity to the TV, words and gestures take on unaccustomed intensity. Anything could happen. The male body, sensing the potential, transduces the heterogeneity of the elements of the situation into a reflex readiness to violence. The “game” is rigged by the male's already-constituted propensity to strike. The typical pattern of relations is re-imposed in the unity of movement of hand against face. The strike expresses the empirical reality of situation: recontainment by the male-dominated power formation of the domestic. The event shortcircuits. The event is recapture. The home event-space is back to the place it was: a container of asymmetric relations between terms already constituted according to gender. Folding back onto domestication. Coded belonging, no becoming”. (Massumi 2002, pp. 80–81).

Non viene descritta una sequenza di azioni, fatti, atti discreti, quanto una trasformazione tutta interna alla superficie di significazione, una trasformazione differenziale che determina un incremento di intensità nello spazio-evento. Potrebbe apparire come una forma di eterogenesi

²⁵⁹ Crenshaw, K. W. (1991). Mapping the margins: Intersectionality, identity Politics, and violence against women of color. *Stanford Law Review* 6: 1241–99.

²⁶⁰ Massumi, B. (2002). *Parables for the virtual: Affect, movement, sensation*. Durham: Duke University Press.

negativamente connotata se vogliamo poiché è il differenziale di intensità tra le diverse posizioni e relazioni dei corpi e oggetti che occupano la superficie dell'evento-casa a modificarsi nella sua intensità.

9.2 La ricerca in impresa: la collaborazione con ART-ER

A partire da questa cornice teorica che ci ha permesso di introdurre brevemente cosa intendiamo per questioni di genere nel contesto della nostra analisi, siamo pronte a restituire una parte del percorso di collaborazione con l'ente che ha cofinanziato questa ricerca. Per accompagnare il lettore e la lettrice in questo passaggio, riteniamo sia necessario specificare che temporalmente il lavoro di ricerca scientifico sopra esposto è avvenuto successivamente alla collaborazione con l'ente cofinanziatore. Questo implica che il piano di analisi del lavoro di ricerca in impresa, come da dottorati PNRR citati in introduzione, presenti un carattere fortemente radicato nel funzionamento e nelle necessità di uno specifico setting o contesto lavorativo, che andiamo a descrivere di seguito.

ART-ER Attrattività Ricerca Territorio è Società Consortile della regione Emilia-Romagna, nata con l'intento di “favorire la crescita sostenibile della regione attraverso lo sviluppo dell'innovazione e della conoscenza, l'attrattività e l'internazionalizzazione del territorio”. Le attività in cui la Società è impegnata sono volte “al rafforzamento dell'ecosistema regionale dell'innovazione e delle relazioni tra i soggetti che lo compongono, alla promozione internazionale del sistema delle imprese e della ricerca, per favorire una crescita sostenibile dei territori e aumentare la capacità di ricerca e innovazione e la loro produttività e a inserirsi con successo nelle catene globali del valore e nei nuovi mercati”.

A strutturare il profilo di ecosistema presiedono diverse aree di intervento, graficamente visibile in Figura 28:

- Europa e Internazionalizzazione
- Investimenti e Costruzioni
- Ricerca e Innovazione
- Sviluppo Sostenibile
- Fondi Strutturali
- Territori Partecipazione Attrattività
- Programmazione Strategica e Studi
- Transizione digitale

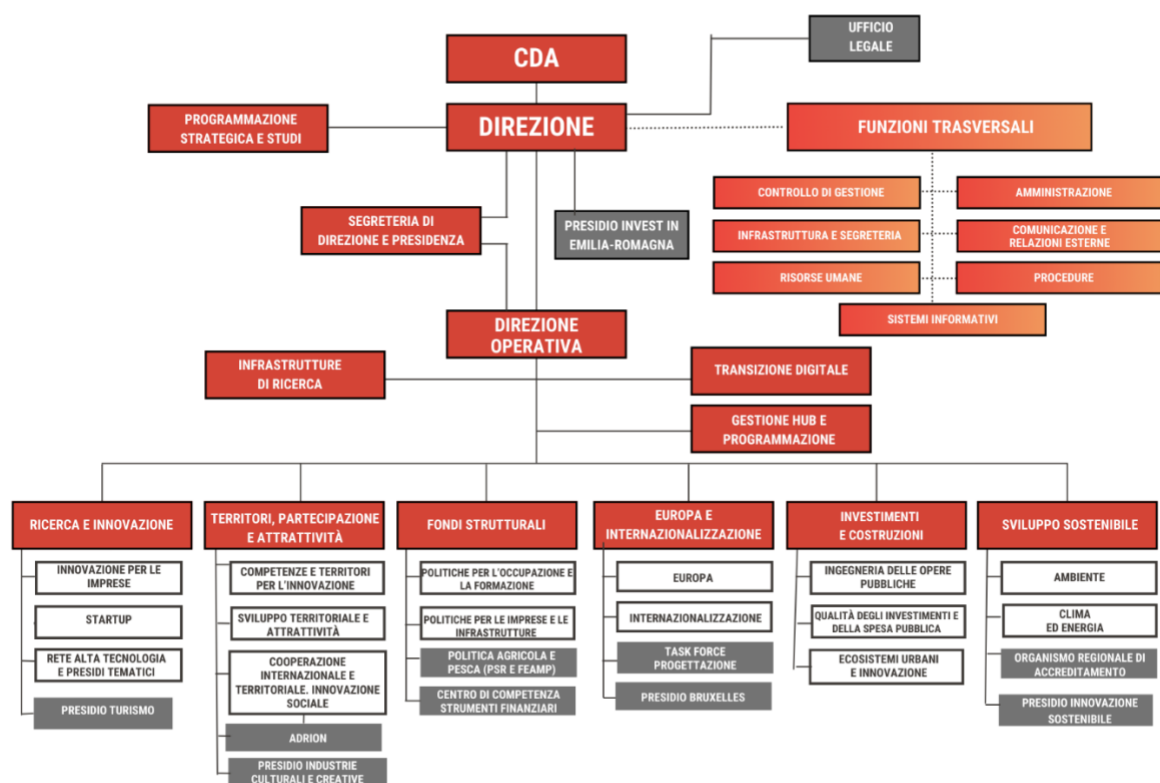


Figura 28: Organigramma e relazioni tra funzioni e comparti.²⁶¹

Il nostro contributo si è inserito nell'ultimo filone di attività la cui responsabile è la dott.ssa Marina Silverii (Direttrice Operativa della Società), ed è stato implementato sotto la guida prima della dott.ssa Rita Trombini, poi del dott. Daniel Caratti. In particolare, l'area Transizione digitale si occupa della pianificazione e dello sviluppo di progettualità nell'ambito dell'Agenda Digitale dell'Emilia Romagna e della diffusione di tecnologie digitali avanzate, tra le quali i sistemi di Intelligenza Artificiale e i Big Data. L'area presiede inoltre alle attività di monitoraggio del territorio dalla prospettiva del processo di digitalizzazione, con un'attenzione particolare al tema delle competenze che intreccia le fasce sociali a rischio esclusione digitale e la cittadinanza tutta. Infatti, l'unità "organizza e coordina azioni di promozione e sviluppo delle competenze digitali per giovani, studenti, insegnanti, adulti e di superamento del gender gap nei percorsi di studio e di lavoro in ambito STEAM".

Il periodo di ricerca presso l'ente cofinanziatore ha coperto 12 mensilità, il più delle quali (circa 10) svolte in avvio del percorso dottorale (annualità 2022/2023) come già anticipato. Nella sua

²⁶¹ Fonte: <https://www.art-er.it/organigramma>

declinazione operativa il progetto è stato incardinato come intervento di valorizzazione e di inquadramento scientifico delle “Azioni Integrate per la diffusione delle competenze digitali in Emilia Romagna” contenute nella delibera della Giunta Regionale con documento n. 1608/2022. A tale fine, è stata sviluppata un’ipotesi di impianto di valutazione dei risultati e dell’impatto delle azioni integrate contenute nella delibera, a integrazione del sistema di monitoraggio in essere e di competenza dell’area di afferenza.

Delle 15 azioni integrate, due sono attualmente oggetto di analisi attraverso l'applicazione sperimentale del modello di valutazione dei risultati e di impatto, secondo un approccio bottom-up:

- A1. Curricoli digitali per adolescenti dell’Emilia-Romagna;
- A2. Competenze digitali per l’apprendimento, il contrasto al gap di genere e povertà educative.

L’approccio bottom-up ha implicato una fase di sperimentazione su singoli interventi verso l’eventuale generalizzazione dello strumento. In riferimento alle suddette azioni, sono state rispettivamente prese in considerazione due progettualità in essere promosse in partnership dalla regione Emilia Romagna. Nel dettaglio:

- per l’Azione 1. il progetto di riferimento è stato il corso di formazione “BIG DATA AND ARTIFICIAL INTELLIGENCE EMILIA-ROMAGNA” alla cui realizzazione hanno collaborato Art-ER, Anpal Servizi, Ufficio Scolastico Regionale, Fondazione Istituto Tecnico Superiore Tecnologie Industrie Creative (FITSTIC), Bi-Rex Big data Innovation & Research Excellence Bologna, Cineca e alcuni partner imprenditoriali fra cui SIS-TER ed Ammagamma srl;
- per l’Azione 2. il riferimento è il progetto regionale “RAGAZZE DIGITALI ER ANNO 2023” PR FSE+.

L’impianto di valutazione da applicare ai due interventi formativi ha presentato queste caratteristiche metodologiche generali: Fase 1: raccolta dati per fotografare il livello di competenze digitali dei partecipanti in chiave autovalutativa, prima o contestualmente all’avvio dell’intervento pubblico. Fase 2: raccolta dati per valutare i risultati, sempre in un’ottica di percezione soggettiva delle competenze acquisite da parte dei/delle partecipanti a conclusione dell’intervento (entro un lasso di tempo ragionevole). Fase 3: raccolta dati di follow-up a distanza di due anni dalla conclusione dell’intervento per valutare l’impatto effettivo dello stesso.



Figura 29: Sintesi visuale impianto di valutazione dei risultati e di impatto.

È importante sottolineare che l'impianto valutativo non ha seguito il modello della valutazione d'impatto di tipo sperimentale che prevede la comparazione tra gruppo soggetto a trattamento e gruppo di controllo, al fine di isolare e determinare il differenziale specifico apportato dall'intervento. Per ragioni legate a estensione potenziale degli interventi coperti dalla Delibera regionale e le contingenze di natura temporale, l'impianto valutativo ha seguito un modello quasi-sperimentale, strutturato sulle comparazioni dei livelli di competenza digitale dei gruppi target nelle tre fasi sopra descritte e le cui variabili da analizzare non sono meccanicamente e causalmente controllate da chi opera la ricerca.

Per quanto concerne gli strumenti di raccolta dati di tipo quantitativo si è proceduto alla costruzione di due questionari in formato digitale somministrati in avvio e a conclusione del corso di formazione. I quesiti posti presentano forma mista strutturata (risposta a scelta singola, multipla, scale Likert) e semistrutturata (risposta aperta, solitamente facoltativa, talvolta condizionata da risposte precedenti).

Sempre a livello metodologico, la valutazione dei risultati prima e di impatto poi ha posto la dimensione di genere come lente di analisi centrale per la valutazione degli obiettivi raggiunti. Indipendentemente dal fatto che il target donne e/o ragazze sia esplicitato come oggetto dell'intervento, l'analisi ha prodotto report con dati disaggregati sulla base del genere in cui i/le partecipanti si riconoscono. Si sceglie espressamente di fare riferimento alla variabile genere e non alla variabile sesso poichè la domanda nel questionario è stata posta nella formula seguente: "In quale genere si/ti riconosce?", lasciando possibilità di risposta aperta e personalizzata alla voce "altro", accanto a Uomo e Donna.

9.2.1 Descrizione primo intervento: Big Data and Artificial Intelligence Emilia-Romagna

Il percorso formativo per docenti, svoltosi tra Marzo e Maggio 2023, è stata una prima sperimentazione di un'attività inquadrata all'interno del progetto "Big Data and Artificial Intelligence Emilia-Romagna" che la Regione ha coordinato in collaborazione con un ampio partenariato di cui fanno parte i soggetti coinvolti nelle attività di progettazione ed erogazione dei moduli.

Gli obiettivi di questa prima sperimentazione sono stati, da un lato, raccogliere elementi per progettare i percorsi PCTO per l'a.s. 2023/2024 per gli/le studenti degli istituti secondari coinvolti nella sperimentazione; dall'altro, valutare gli aspetti positivi e le criticità emerse tanto durante tutte le attività riguardanti il percorso formativo quanto quelle riferite ai PCTO, in funzione di predisporre un modello di intervento replicabile ed estendibile a livello regionale.

Per il nostro specifico contributo, che appunto si è limitato alla costruzione dello strumento di valutazione e all'analisi dei risultati, a livello di progettazione degli strumenti di raccolta dati si è fatto riferimento ad alcuni documenti a carattere europeo e nazionale. Nello specifico, per quanto concerne il questionario in entrata, la definizione di competenze digitali e i loro descrittori e livelli di padronanza sono stati mutuati dal framework [*DigComp Edu*](#)²⁶². In esso sono contenute e organizzate le competenze digitali che il gruppo di lavoro della Commissione Europea ritiene debbano essere possedute dal personale docente ed educativo. Le aree di competenza che compongono il quadro sono:

- **Area 1: Coinvolgimento e valorizzazione professionale**

Usare le tecnologie digitali per la comunicazione organizzativa, la collaborazione e la crescita professionale;

- **Area 2: Risorse digitali**

Individuare, condividere e creare risorse educative digitali;

- **Area 3: Pratiche di insegnamento e apprendimento**

Gestire e organizzare l'utilizzo delle tecnologie digitali nei processi di insegnamento e apprendimento;

- **Area 4: Valutazione dell'apprendimento**

Utilizzare strumenti e strategie digitali per migliorare le pratiche di valutazione;

- **Area 5: Valorizzazione delle potenzialità degli studenti**

²⁶² Pubblicazione originale in inglese dal titolo "*European Framework for the Digital Competence of Educators: DigComp Edu*" Commissione Europea, Joint Research Centre ©European Union, 2017. DOI: 10.2760/159770. ISBN: 978-92-79-73494-6. Versione italiana: ©Istituto per le Tecnologie Didattiche, Consiglio Nazionale delle Ricerche (CNR-ITD).

Utilizzare le tecnologie digitali per favorire una maggiore inclusione, personalizzazione e coinvolgimento attivo degli studenti;

● **Area 6: Favorire lo sviluppo delle competenze digitali degli studenti**

Aiutare gli studenti ad utilizzare in modo creativo e responsabile le tecnologie digitali per attività riguardanti l'informazione, la comunicazione, la creazione di contenuti, il benessere personale e la risoluzione dei problemi.

Ciascuna delle sei aree di competenza è suddivisa a sua volta in una serie di indicatori della competenza associata a una lista di attività che rientrano nelle modalità in cui la competenza in questione si esprime.

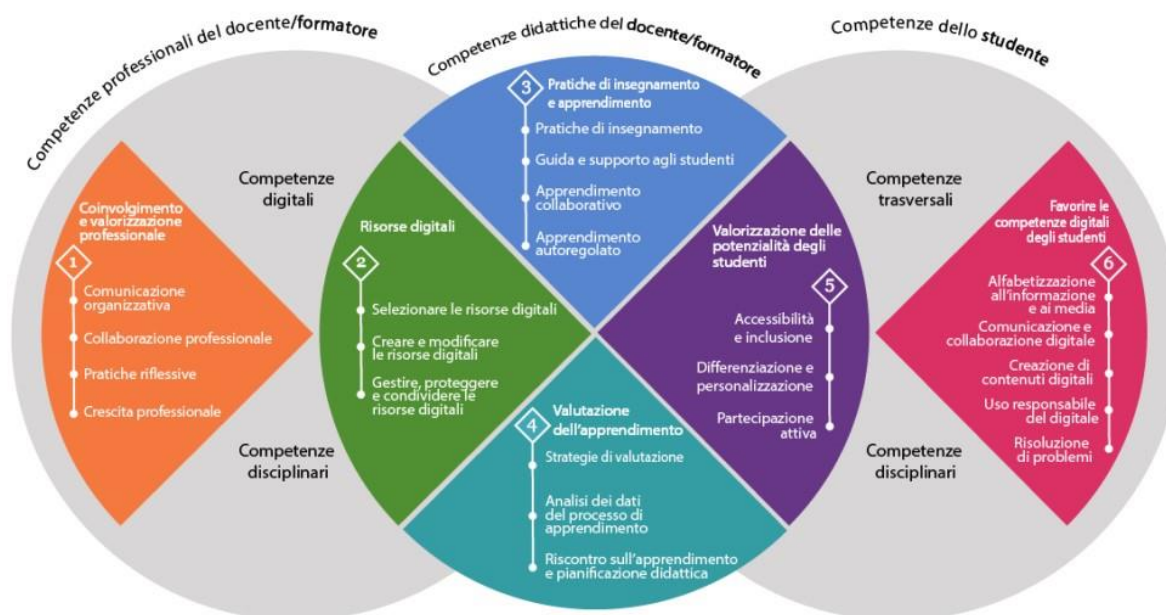


Figura 30: Sintesi del quadro delle competenze digitali del personale docente ed educativo DigComp Edu.

Accanto al DigComp Edu, è stato adottato un ulteriore documento europeo con il titolo di *Orientamenti etici per gli educatori sull'uso dell'intelligenza artificiale (IA) e dei dati nell'insegnamento e nell'apprendimento*²⁶³ non solo in quanto fonte di progettazione degli strumenti di raccolta dati, ma anche come integrazione dello stesso DigComp Edu.

Si legge infatti nei suddetti orientamenti: “Gli educatori e i dirigenti scolastici svolgono un ruolo centrale per il buon esito dell'adozione dei sistemi di IA e per concretizzare i potenziali benefici dei

²⁶³ Commissione europea, Direzione generale dell'Istruzione, della gioventù, dello sport e della cultura, *Orientamenti etici per gli educatori sull'uso dell'intelligenza artificiale (IA) e dei dati nell'insegnamento e nell'apprendimento*, Ufficio delle pubblicazioni dell'Unione europea, 2022, <https://data.europa.eu/doi/10.2766/438>

dati digitali nell'istruzione. Di conseguenza è importante che insegnanti e dirigenti scolastici comprendano e apprezzino le opportunità e le sfide connesse all'impiego dei sistemi di IA, nonché i possibili metodi per potenziare le pratiche di insegnamento, apprendimento e valutazione. *Ne risulterà lo sviluppo di nuove competenze digitali da considerare nel contesto del quadro europeo delle competenze digitali per gli educatori* (DigCompEdu) che costituisce un quadro di riferimento generale a sostegno dello sviluppo in Europa di competenze digitali riguardanti specificamente gli educatori". (CE, 2022: p. 28, corsivo nostro).

Come ultima fonte di progettazione del questionario in ingresso citiamo i risultati della ricerca dal titolo *Big Data and Artificial Intelligence Emilia-Romagna*²⁶⁴ coordinata dalla Regione Emilia-Romagna e realizzata in collaborazione con Anpal, Anpal Servizi spa, ART-ER S. cons. p. a. e in partenariato con università, enti formativi e imprese del territorio. A partire dalla complessa e multifaccettata ricerca frutto del protocollo d'intesa tra Regione Emilia Romagna e Anpal, Anpal Servizi spa, al gruppo docenti rispondenti ai questionari è stato richiesto di autovalutare il proprio livello di competenza relativa alle 7 fasi del Big Data Life Cycle e di stimare una potenziale correlazione tra i profili professionali chiave associati alle fasi del Big Data Life Cycle e il profilo educativo, culturale e professionale (PECUP) del tipo di scuola/indirizzo di studio in cui insegnano. Infine, nel questionario in uscita, accanto a queste fonti di progettazione sono stati aggiunte domande relative ai contenuti didattici oggetto dei moduli formativi e alcuni quesiti relativi al grado di soddisfazione rispetto al corso di formazione nel suo complesso.

Il lavoro di analisi utile a una valutazione dei risultati del percorso formativo è disponibile sul repository DIGIT ER – ART-ER, uno spazio dedicato alla promozione delle attività svolte dalla Società e hub formativo e informativo, con particolare riferimento alle competenze digitali.²⁶⁵

9.2.2 Descrizione secondo intervento: Ragazze Digitali ER 2023

Il secondo intervento a cui è stato applicato l'impianto di valutazione, sia dei risultati che di impatto, ha visto come fascia target una categoria per noi sensibile che è quella di ragazze e giovani donne nel passaggio dalla scuola secondaria di secondo grado ai percorsi di formazione terziaria o ai percorsi lavorativi. In questo caso l'intervento è stato più esteso, coinvolgendo l'intero territorio

²⁶⁴ Direzione Studi & Ricerche – Anpal Servizi, 2022. *Progettazione di un modello di partenariato territoriale tra i sistemi della domanda e dell'istruzione, formazione e lavoro per lo sviluppo delle competenze. Prima sperimentazione in ambito Big Data. Prime conclusioni.* <https://formazioneelavoro.regione.emilia-romagna.it/piani-programmi-progetti/poc-spao/big-data-and-artificial-intelligence-emilia-romagna>

²⁶⁵ digiter.art-er.it

regionale, e presentando cadenza più regolare. Difatti Ragazze Digitali è un’attività che viene proposta dalla Regione in collaborazione con le università emiliano romagnole da diversi anni.

“Ragazze Digitali ER - 2023”, un progetto che ha visto lo svolgimento di 16 Summer Camp gratuiti, a carattere laboratoriale, per studentesse del terzo e quarto anno delle scuole superiori dell’Emilia-Romagna, sui temi del digitale e delle sue applicazioni. Ragazze Digitali ER è inserito nell’ambito delle attività di Data Valley Bene Comune (l’Agenda Digitale dell’Emilia-Romagna) per il quinquennio 2020-2025.

Il target individuato dall’intervento è composto da studentesse frequentanti le scuole secondarie di secondo grado dell’Emilia Romagna delle classi terze e quarte . In continuità con gli obiettivi che accompagnano l’azione 2 della Delibera, la raccolta dati ha indagato il rapporto che le ragazze partecipanti ai Summer Camp hanno con le discipline STEM (Science, Technology, Engineering, Maths) e il loro atteggiamento verso il futuro all’interno di percorsi di studio e lavoro nell’ambito delle tecnologie della comunicazione e dell’informazione, come informatica, elettronica e telecomunicazioni. Come verrà meglio esplicitato nella descrizione del disegno di ricerca che ha preceduto la raccolta dati, obiettivo generale della valutazione della progettualità Ragazze Digitali ER è di verificare quanto tale iniziativa incida nella riduzione del gap di genere nell’accesso ai percorsi di istruzione e formazione post-secondaria e/o a esperienze di lavoro in quegli ambiti disciplinari ancora a prevalenza maschile.

In Politi (2024)²⁶⁶, notavamo, a livello di dati Almalaurea, – l’anno di riferimento in quel caso era il 2022 – “le caratteristiche di un collettivo selezionato comprendente coloro che hanno conseguito la laurea (senza distinzione di livello) nel 2023 nell’area disciplinare STEM in Italia: su 76.948 persone laureate in quest’area, 44.691 si riconoscono nel genere maschile e 32.257 nel genere femminile, con uno scarto di circa 12.400 unità. Andando a disaggregare ulteriormente i dati e isolando il gruppo disciplinare di Informatica e tecnologie ICT otteniamo questo quadro: su un totale di 5.241 persone laureate, 4.342 si riconoscono nel genere maschile e 899 in quello femminile. Così presi questi dati ci dicono forse qualcosa su un problema di rappresentanza, di presenza, di frequenza delle occorrenze, ma non ci dicono nulla da un punto di vista di rappresentazione nella costruzione di sensi e significati” (Ivi, p. 458). Se andiamo a confrontare questi dati con l’anno di riferimento 2024, il gap in ambito ICT resta più o meno immutato: 4.966 laureati si riconoscono nel genere maschile, 1.144 nel genere femminile. A livello di area disciplinare STEM, cresce il numero di “uomini” (46.230), resta per lo

²⁶⁶ Politi, F. (2024). Sull’utilità e sul danno della matematica per la comprensione della semiosi macchinica. *E|C Rivista dell’Associazione Italiana di Studi Semiotici*, 41, 454–463.

più invariato il numero di “donne” (32.780). Alla luce di questo quadro, la Regione Emilia-Romagna promuove diverse iniziative di sensibilizzazione e orientamento, insistendo sul valore aggiunto che una maggiore presenza femminile nell’ambito STEM porterebbe all’economia regionale nel suo complesso. Oltre a ciò, a livello di policy, la stessa delibera 1608/2022 si poneva come “Obiettivo primario della Regione Emilia-Romagna è quindi la diffusione di competenze, responsabilità e consapevolezza digitale a tutti i livelli, dalla formazione all’Intelligenza Artificiale, con un focus specifico rivolto alle fasce marginali, tra cui la popolazione anziana e quelle interessate dal gap di genere” (D.R. 1608/2022, p. 2).

Tornando al nostro contributo, la costruzione degli strumenti di raccolta dati è stata finalizzata alla possibilità di rispondere alla seguente domanda di ricerca: in che misura il progetto Ragazze digitali ha inciso sulla partecipazione (o sul desiderio di partecipazione) di studentesse (genere assegnato) a percorsi formativi o lavorativi post-diploma nell’ambito delle discipline STEM. Tale domanda ha reso necessario definire una serie di aspetti che devono essere meglio chiariti e specificati:

- quali percorsi post-diploma prendere in considerazione: università, istituti tecnici superiori, percorsi di formazione specializzata, inserimento lavorativo;
- definizione precisa degli ambiti disciplinari (qualsiasi ambito afferente alle discipline STEM oppure focus su SSD informatica);
- ammissibilità di una misurazione degli atteggiamenti verso materie tecnico-scientifiche a scuola nei mesi/anni successivi alla partecipazione al corso di formazione.

Entrando nel dettaglio delle dimensioni da analizzare attraverso il questionario in ingresso, sono state prese in considerazione:

- Le informazioni anagrafiche: paese di nascita (se diverso da Italia, da quanti anni in Italia), provincia di residenza, comune di residenza, genere in cui ci si riconosce, anni, classe frequentata, scuola frequentata (tipo e indirizzo);
- Le aspettative formative e di orientamento con riferimento teorico alla Self-Determination Theory²⁶⁷: analisi motivazionale [intrinseca, estrinseca, amotivazionale] della partecipazione al percorso formativo, interesse per materie STEM, proiezione di sé nel futuro (sia a livello di percorsi di istruzione e formazione, sia a livello lavorativo), azione del role modelling (familiare o sociale),

²⁶⁷ Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation development and wellness*. New York, NY: Guilford Press; Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, Article 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>

percezione della rilevanza/prossimità degli argomenti trattati nei corsi, gradi di autonomia rispetto agli argomenti trattati nei corsi;

- Conoscenze e competenze pregresse: percezione di autoefficacia rispetto ai risultati scolastici, ai contenuti specifici dei corsi, ai descrittori di competenze contenuti nel quadro DigComp 2.2, contesti di apprendimento/avvicinamento a discipline STEM;
- Aspetti organizzativi dei corsi e loro rappresentazione come facilitatori oppure ostacoli alla partecipazione: presenza di sole ragazze (genere assegnato), scansione oraria degli incontri, durata complessiva, sedi di svolgimento, materiali e attrezzatura a disposizione, tipologia di argomenti affrontati.

Per quanto concerne invece le dimensioni da analizzare attraverso il questionario in uscita, si è proceduto ad una costruzione paratattica rispetto a quelle presentate nel questionario in ingresso, aggiungendo ciò che non era di principio misurabile in avvio di percorso:

- Informazioni di tipo anagrafico.
- Grado di soddisfazione in relazione alle aspettative formative e di orientamento.
- Conoscenze e competenze acquisite.
- Mutato atteggiamento nei confronti delle discipline STEM.
- Potenziale di empowerment di alcuni fattori del percorso formativo: argomenti e tematiche oggetto della formazione, informalità del contesto di apprendimento, composizione del gruppo.
- Grado di soddisfazione in relazione ad aspetti organizzativi - caratteristiche dei percorsi: come per il questionario in ingresso, con l'aggiunta delle modalità di erogazione dei contenuti.
- Orientamenti per il futuro.

Nella progettazione degli strumenti di raccolta dati si è fatto riferimento ad alcune fonti a carattere accademico e istituzionale.

Nello specifico, per quanto concerne il questionario in entrata e in uscita, la definizione di competenze digitali e i loro descrittori e livelli di padronanza sono stati mutuati dal framework DigComp 2.2.²⁶⁸ In esso sono contenute e organizzate le competenze digitali che il gruppo di lavoro

²⁶⁸ Vuorikari, R., Kluzer, S. and Punie, Y., *DigComp 2.2: The Digital Competence Framework for Citizens*, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-48882-8. <https://publications.jrc.ec.europa.eu/repository/handle/JRC128415>

Versione italiana *Il Quadro delle Competenze Digitali per i Cittadini - Versione 2.2*. Gruppo di Lavoro coordinato da Pasquale Popolizio, per il Dipartimento per la trasformazione digitale della Presidenza del Consiglio dei ministri, e co-coordinato da Sandra Troia e Stefano Kluzer. Dicembre 2022

della Commissione Europea ritiene debbano essere possedute da cittadini e cittadine per una sana partecipazione alla vita sociale. Le aree di competenza che compongono il quadro sono:

- **Alfabetizzazione su informazione e dati**

Articolare le esigenze informative, individuare e recuperare dati, informazioni e contenuti digitali. Giudicare la rilevanza della fonte e del suo contenuto. Archiviare, gestire e organizzare dati, informazioni e contenuti digitali.

- **Comunicazione e collaborazione**

Interagire, comunicare e collaborare tramite le tecnologie digitali, tenendo conto della diversità culturale e generazionale. Partecipare alla società attraverso i servizi digitali pubblici e privati e la cittadinanza attiva. Gestire la propria presenza, identità e reputazione digitale.

- **Creazione di contenuti digitali**

Creare e modificare contenuti digitali. Migliorare e integrare le informazioni e i contenuti in un corpus di conoscenze esistenti, comprendendo come applicare il copyright e le licenze. Saper dare istruzioni comprensibili ad un sistema informatico

- **Sicurezza**

Proteggere i dispositivi, i contenuti, i dati personali e la privacy negli ambienti digitali. Proteggere la salute fisica e psicologica ed essere competenti in materia di tecnologie digitali per il benessere e l'inclusione sociale. Essere consapevoli dell'impatto ambientale delle tecnologie digitali e del loro utilizzo.

- **Risolvere problemi**

Identificare esigenze e problemi e risolvere difficoltà concettuali e situazioni problematiche in ambienti digitali. Utilizzare gli strumenti digitali per innovare processi e prodotti. Mantenersi aggiornati sull'evoluzione del digitale.

Ciascuna delle cinque aree di competenza è suddivisa a sua volta in una serie di indicatori della competenza associata ai livelli di padronanza, a esempi di conoscenze, attività e attitudini e a casi d'uso in contesto formativo e lavorativo.

Accanto al DigComp 2.2 per la definizione e l'inquadramento delle competenze digitali, è stata utilizzata la letteratura scientifica relativa al rapporto tra genere e discipline STEM. Si citano in proposito:

- Biemmi I., Leonelli S. (2016). *Gabbie di genere. Retaggi sessisti e scelte formative*. Rosenberg & Sellier;

- Hermans, S., Gijzen, M., Mombaers, T., & van Petegem, P. (2022). Gendered patterns in students' motivation profiles regarding iSTEM and STEM test scores: A cluster analysis. *International Journal of STEM Education*, 9(1).

utilizzati rispettivamente per la parte di questionario relativa ai condizionamenti impliciti ed espliciti agiti dai soggetti di orientamento (famiglia, scuola, società) e per la parte dedicata all'indagine motivazionale della partecipazione ai summer camp.



Figura 31: Sintesi del quadro delle competenze digitali per la cittadinanza DigComp 2.2.

Nella fase di follow up, ossia a due anni dalla conclusione dei summer camp, a livello di strumenti di raccolta dati si è proceduto alla costruzione di un terzo questionario, strutturalmente più snello e di rapida compilazione, che è stato somministrato digitalmente alle rispondenti. I quesiti posti presentano forma mista: risposta a scelta multipla e risposta aperta, condizionate da risposte precedenti.

La costruzione degli strumenti di raccolta dati è stata finalizzata alla possibilità di rispondere alla seguente domanda di ricerca: in che misura il progetto Ragazze digitali ha inciso sulla partecipazione (o sul desiderio di partecipazione) di studentesse (genere assegnato) a percorsi formativi o lavorativi post-diploma nell'ambito delle discipline STEM.

Nel dettaglio delle dimensioni analizzate attraverso il questionario di follow up, se ne descrive la struttura di seguito:

- Le informazioni di tipo anagrafico: nome e cognome, indirizzo email, età, summer camp frequentato, eventuale variazione di scuola o indirizzo successiva alla partecipazione ai summer camp.
- Scelta del percorso post-diploma, già compiuta o in fase di compimento, con domanda di specificazione della scelta sulla base della selezione precedente: le alternative proposte sono state:
 - Iscrizione ad un corso di laurea in ambito STEM (Scientifico, Architettura e Ingegneria civile, Ingegneria industriale e dell'informazione)
 - Iscrizione ad un corso di laurea in ambito Informatica e Tecnologie ICT (tecnologie della comunicazione e dell'informazione)
 - Iscrizione ad un corso di laurea non STEM
 - Iscrizione a un Istituto Tecnico Superiore
 - Ingresso nel mondo del lavoro nell'ambito delle discipline STEM
 - Ingresso nel mondo del lavoro in contesti non legati alle discipline STEM
 - Indecisione
 - Altro

Per ciascuna delle alternative, seguono delle ulteriori domande di specificazione chiarimento, al fine di inquadrare meglio la scelta.

- Due domande a risposta aperte per un primo inquadramento del percorso che ha portato alla scelta post-diploma e della interpretazione soggettiva del concetto di competenza digitale. In questa sezione è stato inoltre chiesto di indicare la modalità espressiva con la quale le partecipanti si sentirebbero maggiormente a loro agio nel raccontare la propria esperienza di scelta post-diploma.

I report prodotti al termine dei Summer Camp e a distanza di due anni dalla conclusione degli stessi, per la parte di follow-up, sono stati redatti per un uso interno alle attività di collaborazione con l'ente cofinanziatore.

Apriamo, in conclusione, una parentesi che interessa il piano metodologico: la valutazione d'impatto contribuisce alla messa in luce della correlazione tra un effetto e la sua causa²⁶⁹, isolando porzioni di senso da un processo evidentemente più complesso. Tuttavia, occorre sottolineare come la scelta dello strumento di analisi sia funzionale ad una certa configurazione material-discorsiva (riprendiamo qui il discorso di Barad sugli apparati), ma anche il nodo del rapporto tra pratiche di riflessività e ricerca quantitativa risuona nella scelta di tali strumenti.

Nella logica del genere come spazio-evento sopra descritta, i fattori che influenzano la scelta post-diploma possono essere analizzati come eventi significanti o forme del senso emergenti all'interno di una dinamica non strutturalmente predeterminata (come i percorsi di maturazione di una scelta/decisione), in uno specifico contesto storico-sociale con i suoi impatti sulla definizione di ciò che ha valore e sulle scale di priorità. Da questa prospettiva la scelta degli strumenti e delle pratiche di analisi che sono stati adottati, per ragioni contingenti all'organizzazione del lavoro di ricerca in impresa, potrebbe essere successivamente integrata andando a scandagliare, ad esempio, come le forze e le leve che compongono la scelta di un percorso post-diploma cambino a seguito di una serie di interventi a livello di policy, studiando quali atteggiamenti, convinzioni, abiti si magnificano e quali si narcotizzano nelle rappresentazioni testuali. La semiotica, in questo, può essere un'alleata importante del decisore politico, indagando le trasformazioni del senso all'interno di una cultura e delle sue istituzioni.

²⁶⁹ Per una rassegna teorica si vedano: Checchi, Trivellato (2014). *La valutazione d'impatto delle politiche in Italia tra sviluppo dei metodi e affermazione di buone pratiche: un'introduzione*; Cortoni, Lo Presti (2014). *Verso un modello di valutazione delle competenze digitali*; Vecchi, G. (2006). *Il monitoraggio e la valutazione intermedia di programmi complessi*; Martini, A. (2006). *Metodo sperimentale, approccio controfattuale e valutazione degli effetti delle politiche pubbliche*.

10. A guisa di conclusione

Siamo pronte, a chiusura di questo lavoro di ricerca, a trarre delle conclusioni che non si limitino a ripercorrere quanto esposto nei capitoli precedenti, ma che sappiano individuare *i punti critici*, oseremmo dire, cioè i punti di svolta del nostro testo. Proviamo quindi a mettere in luce leve positive e negative, così da giungere fluidamente verso potenziali futuri terreni di ricerca.

Partiamo dal “metodo di lavoro”. In introduzione abbiamo anticipato che il procedere per analogia di proporzionalità avrebbe costituito la nostra base di metodo. Da tale prospettiva applicata, questo elaborato potrebbe non rispettare i caratteri standard del metodo semiotico applicato alla testualità chiusa, coesa e coerente: il corpus è frammentato, l’analisi manca di organicità terminologica e disciplinare e percorre lunghe distanze nella considerazione dei suoi oggetti. Se da un lato questo può presentarsi come un limite, dall’altro ci è sembrato tuttavia euristico nel suo voler ricostruire percorrenze o cammini inferenziali tra domini le cui interiorità epistemiche sono di difficile accesso a chi non è del mestiere. Il nostro progetto ha usato una teoria della traduzione per effettuare traduzioni tra modelli epistemologici. L’accortezza, di cui abbiamo tenuto conto lungo tutta la stesura, è stata di non cedere alla seduzione del sistema dei saperi e delle conoscenze; ancora una volta il nostro bersaglio polemico è la dialettica hegeliana. Non c’è una gerarchia di saperi e conoscenze, ci sono tante prospettive parziali che illuminano un oggetto “sotto un certo rispetto o capacità”. In questo consiste il primato epistemologico del locale sul globale e, laddove invece si rivendichi l’uno, faranno sempre da contraltare molteplicità di esclusioni. Far dialogare tra loro campi di studio e metodi di ricerca non ha significato per noi ridurre a sintesi i rispettivi modi di fare scienza; ma questo dialogo ha tentato di attraversare porzioni locali di realtà o piani enciclopedici che in alcuni casi sono inconsapevoli delle reciproche implicazioni. Riteniamo che la Semiotica sia propriamente quel campo che mette a disposizione gli strumenti teorici necessari a operare questi passaggi di mediazione e traduzione tra circuiti di significazione. Questa scelta di metodo si è configurata, dunque, come un posizionamento epistemologico e come una dichiarazione di intenti.

Sulla validità dell’ipotesi teorica, ossia sull’idea che il rapporto tra prompt e modello di linguaggio funzioni plasticamente come l’attivazione dell’Enciclopedia di Eco per il tramite dell’atto di enunciazione, riteniamo che non possa spettare a noi valutarne la bontà. Guardiamo al mondo “sedute sulle spalle di giganti” e a loro non possiamo che affidare con riconoscenza queste poche riflessioni tentative. Riteniamo che le formazioni concettuali dell’enunciazione impersonale ed evenemenziale, dell’enciclopedia, del virtuale intensivo, dell’eterogenesi differenziale, della sintassi strutturale e dello schematismo posizionale morfodinamico siano potenti dispositivi narrativi della realtà e le connessioni tra loro rendano “tutto luminoso”, al netto dei nostri *confirmation bias*.

Un'ultima considerazione sulle questioni di genere o sul genere come questione: pensiamo che la Semiotica possa assumere la *vulnerabilità* come modo proprio di fare scienza, un tratto specifico che tiene insieme l'elevazione a minore e la complessità. Un'ecologia semiotica della complessità risponde alla tentazione di unificazione e messa a sistema di ipotesi interpretative che sono sempre e costitutivamente locali. Ci piacerebbe che il fare semiotico, inteso come sguardo e pensiero semiotico, colga la potenza delle marginalità e con attenzione e cura si impegni nella ricostruzione delle sue più o meno visibili connessioni.

Poste queste tre brevi premesse, mettiamo a fuoco per l'appunto le connessioni non sviluppate – per quelle non individuate, siamo aperte a nuove esplorazioni – che presentiamo come tracce di potenziali sviluppi di ricerca futuri.

- i) In primo luogo, vorremmo ritornare sulla tesi centrale, citata in premessa: sarebbe per noi auspicabile tentarne una validazione operativa e di messa a terra. Intendiamo cioè che sarebbe importante approfondire lo stato dell'arte sull'addestramento dei LLM, all'intersezione tra linguistica computazionale e machine learning, per testare un'ipotesi di preparazione dei dati basata sulla linea teorica Tesnière, Peirce, Petitot, mantenendo in background gli apporti teorici forniti dalla linea Eco, Paolucci, Sarti. L'obiettivo vorrebbe essere quello di ridurre il quantitativo di dati necessari all'addestramento dei modelli di linguaggio, valutandone l'efficacia non su grandi architetture, ma su architetture di scala ridotta. Potremmo definirlo un movimento di *scaling down*, alternativo all'ipotesi di *scaling up* a sostegno delle proprietà emergenti dei modelli di linguaggio.
- ii) Una seconda pista d'indagine che ci sembra aver lasciato aperta riguarda la generazione del testo. Abbiamo visto come, nell'essere umano, i processi linguistici e cognitivi superiori implicati nella produzione e pianificazione di un testo scritto siano molteplici e complessi. La narrazione prevalente sulla produzione del testo da parte dei grandi modelli di linguaggio si limita a dire che la generazione dell'output procede per probabilità e predizione del token successivo. Ancora una volta una parzialità che si propone come totalizzante. La tenuta concettuale di questa affermazione ci appare limitata e riteniamo sarebbe interessante sviluppare un lavoro di analisi che tenga insieme i processi computazionali necessari alla produzione di un testo da parte delle macchine dotate di linguaggio e i processi cognitivi e linguistici che presiedono alla stessa operazione.
- iii) Da una prospettiva più strettamente teorica, ci sembra vada ulteriormente approfondito il concetto di virtualità alla luce del funzionamento dei modelli di linguaggio. Abbiamo tracciato alcune possibili linee di percorrenza che da Saussure a Eco, passando per

Dondero, D'Armenio, Paolucci, Sarti, si collegano a Deleuze e Foucault. Qui centrale è la relazione tra enunciazione e archivio; riteniamo andrebbe indagato meglio il piano della composizione delle forze e come quest'ultime influenzino i passaggi tra modi di esistenza all'interno dell'enunciato semiotico. Il dialogo con le matematiche da questo punto di vista ci sembra particolarmente fruttuoso.

- iv) Infine, le policy per la diffusione di competenze digitali. Il nostro contributo nella collaborazione con ART-ER è stato sperimentale e meriterebbe un lavoro di sistematizzazione, oltre che di approfondimento. Alla necessaria operatività richiesta in fase di avvio del percorso dottorale dovrebbe corrispondere, in conclusione dello stesso, un apporto più riflessivo e calibrato su bisogni specifici e criteri di sostenibilità, efficacia e riproducibilità. Abbiamo intravisto queste esigenze e riteniamo sia possibile darne risposta.

Per concludere, siamo consapevoli della parzialità delle argomentazioni proposte, ma a nostra difesa ci piace pensare che la ricerca, il fare ricerca, abbia proprio questo tratto da valorizzare: una potenziale infinitezza e una costante perfettibilità asintotica.

Bibliografia

Al-Rfou, R., Choe, D., Constant, N., Guo, M., & Jones, L. (2018). *Character-Level Language Modeling with Deeper Self-Attention*. arXiv:1808.04444. <https://doi.org/10.48550/arXiv.1808.04444>

Alemohammad, S., Humayun, A. I., Agarwal, S., Collomosse, J., & Baraniuk, R. (2024). *Self-Improving Diffusion Models with Synthetic Data* (No. arXiv:2408.16333). arXiv. <https://doi.org/10.48550/arXiv.2408.16333>

Alipour, S., Galeazzi, A., Sangiorgio, E., Avalle, M., Bojic, L., Cinelli, M., & Quattrociocchi, W. (2024). Cross-platform social dynamics: An analysis of ChatGPT and COVID-19 vaccine conversations. *Scientific Reports*, 14(1), 2789.

Ambridge, B. & Blything, L. (2024). Large language models are better than theoretical linguists at theoretical linguistics. *Theoretical Linguistics*, 50(1-2), 33-48. <https://doi-org.ezproxy.unibo.it/10.1515/tl-2024-2002>

Anderson, C. (2025). *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*. Wired. Recuperato 25 settembre 2025, da <https://www.wired.com/2008/06/pb-theory/>.

Anderson, G., Verd, B., & Jaeger, J. (2020). Drawing to Extend Waddington's Epigenetic Landscape. *Leonardo*, 53(3), 256–262. https://doi.org/10.1162/leon_a_01738

Aristotele, *Topici* in Colli, G. (1955). Organon. Einaudi.

Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society* 28, no. 3: 801–831.

Barad, K. (2007). *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. In *Meeting the Universe Halfway*. Duke University Press. <https://doi.org/10.1515/9780822388128>

Basso Fossali, P. (2017). *Vers une écologie sémiotique de la culture. Perception, gestion et réappropriation du sens*. Limoges: Lambert-Lucas.

Basso Fossali, P. (2007). Di mediazione in mediazione. Spazi esperienziali, domini culturali e semiosfera. *Visible*, (3). <https://doi.org/10.25965/visible.197>

Basso Fossali, P. (2025). Rationalités correctives et intelligence artificielle assistée: les doubles contraintes des humanités numériques. *Semiotica*, 2025(262), 71–109. <https://doi.org/10.1515/sem-2024-0189>

Basso Fossali, P. (2025a), *Forme di vita e scenarizzazioni dell'indisponibile. Per un'ecologia semiotica della proporzione*, in XXX Congresso della Società di Filosofia del Linguaggio “I linguaggi della crisi. Clima, Conflitto, Antropocene”.

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.463>

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). <https://doi.org/10.1145/3442188.3445922>

Bennett L. (2023). A Biosemiotic Approach to AI. Folding as semiotic modeling. Translated by A. Lyubimova, *Method* 3(4), pp. 147173.

Bertoni, F., Montobbio, N., Sarti, A., & Citti, G. (2021). *Emergence of Lie symmetries in functional architectures learned by CNNs* (No. arXiv:2104.08537). arXiv. <http://arxiv.org/abs/2104.08537>

Biemmi I., Leonelli S. (2016). *Gabbie di genere. Retaggi sessisti e scelte formative*. Rosenberg & Sellier.

Brown, t. et al. (2020). Language models are few-shot learners. In *Advances in neural information processing systems*, vol. 33, pp. 1877–1901.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91.

Butler, J. (1999). *Gender trouble: Feminism and the subversion of identity* (10th anniversary ed.). Routledge. — Trad. it.: Butler, J. (2017). *Questione di genere. Il femminismo e la sovversione dell'identità* (S. Adamo, Trad.). Laterza.

Capone, L., Suozzi, A., Lebani, G. E., & Lenci, A. (2025). *BAMBI Goes to School: Evaluating Italian BabyLMs with Invalsi-ITA*. CLiC-it 2025: Eleventh Italian Conference on Computational Linguistics, Cagliari.

Capone, L. (2024) al XXIX Congresso della Società di Filosofia del linguaggio “Filosofia dei linguaggi e paradigmi psicologici” dal titolo: “*Un'analisi psicolinguistica delle strategie di apprendimento dei Language Model*”

Capone, L. (Forthcoming). Child Development and Language Model Training: Comparing Psychological Theories of Learning. *RIFL SFL-Filosofia dei linguaggi e paradigmi psicologici*.

Cardon, D., Cointet, J.-P., & Mazières, A. (2018). La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, 211(5), 173–220. <https://doi.org/10.3917/res.211.0173>

Caronia, A. (1986), L'uomo artificiale, in id. (2020), *Dal cyborg al postumano. Biopolitica dell'uomo artificiale*. Meltemi.

Chersoni, E. M. (2021). *Distributional Semantics and Event Knowledge* [slide deck]. Università di Hildesheim.

Chollet, F. (2021). *Deep Learning with Python*, Second Edition. Manning Publications Co. LLC.

Chu, Z., Chen, J., Chen, Q., Yu, W., He, T., Wang, H., Peng, W., Liu, M., Qin, B., & Liu, T. (2024). Navigate through Enigmatic Labyrinth A Survey of Chain of Thought Reasoning: Advances, Frontiers and Future. In L.-W. Ku, A. Martins, & V. Srikumar (A c. Di), *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers) (pp. 1173–1203). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.acl-long.65>

Clark, A. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. Oxford University Press.

Clark, A. (2023). *The Experience Machine: How Our Minds Predict and Shape Reality*. Random House.

Clark, A. (2025). Extending Minds with Generative AI. *Nature Communications*, 16(1), 4627. <https://doi.org/10.1038/s41467-025-59906-9>.

Cocchi, F., Moratelli, N., Cornia, M., Baraldi, L., & Cucchiara, R. (2025). *Augmenting Multimodal LLMs with Self-Reflective Tokens for Knowledge-based Visual Question Answering*. In IEEE/CVF Conference on Computer Vision and Pattern Recognition Nashville TN June 11th - June 15th.

Coeckelbergh, M. (2021). Time Machines: Artificial Intelligence, Process, and Narrative. *Philosophy & Technology*, 34(4), 1623–1638. <https://doi.org/10.1007/s13347-021-00479-y>.

Colas-Blaise, M. (2025). La machine crée, mais énonce-t-elle? Le computationnel et le digital mis en débat. *Semiotica*, 2025(262), 147–187. <https://doi.org/10.1515/sem-2024-0188>

Commissione europea. Direzione generale dell'Istruzione, della gioventù, dello sport e della cultura. (2022). *Orientamenti etici per gli educatori sull'uso dell'intelligenza artificiale (IA) e dei dati nell'insegnamento e nell'apprendimento*. Ufficio delle pubblicazioni dell'Unione europea. <https://data.europa.eu/doi/10.2766/438>

Crenshaw, K. W. (1991). Mapping the margins: Intersectionality, identity Politics, and violence against women of color. *Stanford Law Review* 6: 1241–99.

Cristianini, N. (2024). *Machina Sapiens. L'algoritmo che ci ha rubato il segreto della conoscenza*. Bologna: Il Mulino.

Cristianini, N. (2025). *Sovrumano. Oltre i limiti della nostra intelligenza*. Bologna: Il Mulino.

Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press; tr. It. *Né intelligente né artificiale. Il lato oscuro dell'IA*. Bologna: Il Mulino (2021).

Currier, Dianne (2003). Feminist technological futures: Deleuze and body/technology assemblages. *Feminist Theory* 3: 321–38

D'Armenio, E. (2025). Les IA génératives visuelles entre perception d'archives et circuits de composition. *Semiotica*, 2025(262), 213–257. <https://doi.org/10.1515/sem-2024-0184>

D'Armenio, E., Dondero, M. G., Deliège, A., & Sarti, A. (2025). For a Semiotic Approach to Generative Image AI. *Semiotic Review*, 9. <https://doi.org/10.71743/EE5NRX33>

D'Armenio, E., Rosso, A., & Voto, C. (2025), Sémiotique de l'espace à l'ère de la logique computationnelle et des pratiques numériques. *Signata. Annales des sémiotiques / Annals of Semiotics*, 16. <https://doi.org/10.4000/14rhb>.

Deliège, A. (2024). *Énonciation(s) et passions dans les territoires sémiotiques ouverts par l'Intelligence Artificielle* in Séminaire International de Sémiotique à Paris 2023-2024: <https://www.youtube.com/watch?v=sGmgQUGF4Z8>

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41(6), 391–407.

Degli Esposti, P., Bonazzi, M., & Spillare, S. (2024). AI Imaginaries and Narratives in the Italian Public Discourse: The Impact of ChatGPT. *IM@GO*, 23, 132–147.

Deleuze, G. (1966). Review of Gilbert Simondon's *L'individu et sa genèse physico-biologique*. *Pli*, 12(2001), 43-49.

Deleuze, G. (1968). *Différence et répétition*, Paris: Presses Universitaires de France; tr. it. *Differenza e ripetizione*, Il Mulino (1971).

Deleuze, G. (1986). *Foucault*. Les Éditions de Minuit; tr. it *Foucault*. Orthotes (2018).

Deleuze, G., & Guattari, F. (1980). *Mille plateaux. Capitalisme et schizophrénie*. Les Éditions de Minuit. — Trad. it.: Deleuze, G., & Guattari, F. (2024). *Mille piani. Capitalismo e schizofrenia 2*. Orthotes.

Demaria, C. (2016). Intersezionalità e femminismo transnazionale tra costruttivismo, post-strutturalismo e 'performance' epistemologiche. *Scienza & Politica. Per una storia delle dottrine*, Vol. 28, No 54 (2016). <https://doi.org/10.6092/ISSN.1825-9618/6219>

Demaria, C. (2019). *Teorie di genere. Femminismi e semiotica*: Milano: Bompiani.

Dondero, M. G. (2020). *Les langages de l'image. De la peinture aux Big Visual*. Paris: Hermann Éditeurs; tr. it. *I linguaggi dell'immagine. Dalla pittura ai Big Visual Data*. Milano: Meltemi.

Dondero, M. G. (2025). Semiotics of artificial intelligence: enunciative praxis in image analysis and generation. *Semiotica*. 2025(262), 111–146. <https://doi.org/10.1515/sem-2024-0195>

Dondero, M., Alonso Aldama, J., & Leone, M. (2025). Aspects of AI semiotics: enunciation, agency, and creativity. *Semiotica*, 2025(262), 1–3. <https://doi.org/10.1515/sem-2025-0014>

Dong, Q., Li, L., Dai, D., Zheng, C., Ma, J., Li, R., Xia, H., Xu, J., Wu, Z., Chang, B., Sun, X., Li, L., & Sui, Z. (2024). A Survey on In-context Learning. In Al-Onaizan, Y., Bansal, M. & Chen, Y. (A c. Di), *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 1107–1128). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.emnlp-main.64>

Drage, E., & Frabetti, F. (2023). *AI that Matters: A Feminist Approach to the Study of Intelligent Machines*. In J. Browne, S. Cave, E. Drage, & K. McInerney (A c. Di), *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machine*. Oxford University Press. <https://doi.org/10.1093/oso/9780192889898.003.0016>

Duarte, F. (2025). Number of ChatGPT Users (July 2025). *Exploding Topics*. <https://explodingtopics.com/blog/chatgpt-users#how-many>

Dusi, N., (2000), “Introduzione. Per una ridefinizione della traduzione intersemiotica”, *VERSUS. Quaderni di studi semiotici*, monografico: La traduzione intersemiotica, n. 85-86-87, pp. 3-54

DWF (2024) *Femministe col Bot: Tecnologie e Intelligenze Artificiali*. DWF (142) 2024, 2 | <https://www.dwf.it/rivista/femministe-col-bot-dwf-142-2024-2/>

Eco, U. (1964). *Apocalittici e integrati*. Milano: Bompiani. (Edizione digitale 2017).

Eco, U. (1968). *La struttura assente*. Milano: Bompiani. (La Nave di Teseo 2016).

Eco, U. (1975). *Trattato di semiotica generale*. Milano: Bompiani. (La Nave di Teseo 2016).

Eco, U. (1979). *Lector in fabula*. Milano: Bompiani. (La nave di Teseo ed. 2020).

Eco, U. (1983), *L'Antiporfirio*, in *Il pensiero debole* a cura di Vattimo G. e Rovatti P.A. Feltrinelli. Ripubblicato in Eco, U. (1994).

Eco, U. (1984). *Semiotica e filosofia del linguaggio*. Torino: Einaudi.

Eco, U. (1990). *I limiti dell'interpretazione*. Milano: Bompiani.

Eco, U. (1992). *The original and the copy*. In F. J. Varela & J.-P. Dupuy (Eds.), *Understanding origins*. Dordrecht: Kluwer.

Eco, U. (1994), *Sugli specchi e altri saggi. Il segno, la rappresentazione, l'illusione, l'immagine*. (La nave di Teseo ed. 2018).

Eco, U. (1997). *Kant e l'ornitorinco*. Milano: Bompiani.

Eco, U. (2004). 4. L'analogia entis. *Doctor Virtualis*, (3). <https://doi.org/10.13130/2035-7362/55>

Eco, U. (2007). *Dall'albero al labirinto. Studi storici sul segno e l'interpretazione*. La nave di Teseo.

Eco, U. (2012). *Dire quasi la stessa cosa. Esperienze di traduzione*. Milano: Bompiani.

Elman J. L. (2011). Lexical knowledge without a lexicon?. *The mental lexicon*, 6(1), 1–33. <https://doi.org/10.1075/ml.6.1.01elm>

Evans, O., Cotton-Barratt, O., Finnveden, L., Bales, A., Balwit, A., Wills, P., Righetti, L., & Saunders, W. (2021). *Truthful AI: Developing and governing AI that does not lie* (No. arXiv:2110.06674). arXiv. <http://arxiv.org/abs/2110.06674>

Fabbri, P. (1998). *La svolta semiotica*. Roma-Bari: Laterza.

Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L., & Agarwal, S. (2025). *How AI and Human Behaviors Shape Psychosocial Effects of Chatbot Use: A Longitudinal Randomized Controlled Study*. MIT Media Lab (preprint)

Fazi, M. B. (2024). Machines That Create: Contingent Computation and Generative AI. *Media Theory*, 8(2), Art. 2. <https://doi.org/10.70064/mt.v8i2.1115>

Ferraris, M. (2021). *Documanità [Documanity]*. Laterza.

Ferraris, M. (2025). *La pelle. Che cosa significa pensare nell'epoca dell'Intelligenza Artificiale*. Il Mulino.

Firth, John R. 1957. A synopsis of linguistic theory 1930–1955. Pages 1–32 of: *Studies in Linguistic Analysis*. Oxford: Blackwell

Floridi, L. (2022). *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*. Raffello Cortina Editore.

Floridi, L. (2025a). *La differenza fondamentale. Artificial Agency: una nuova filosofia dell'intelligenza artificiale*. Mondadori.

Floridi, L. (2025b). AI as Agency without Intelligence: On Artificial Intelligence as a New Form of Artificial Agency and the Multiple Realisability of Agency Thesis. *Philosophy & Technology*, 38(1), 1–27.

Floridi, L., Buttaboni, C., Hine, E., Morley, J., Novelli, C., & Schroder, T. (2025). Agentic AI Optimisation (AAIO): What it is, how it works, why it matters, and how to deal with it (arXiv:2504.12482). arXiv. <https://doi.org/10.48550/arXiv.2504.12482>

Fontanille, J. (2003). *Sémiotique du discours*. Pulim.

Fontanille, J. (2004), *Soma et séma. Figures du corps*, Maisonneuve et Larose, Paris; tr. it. *Figure del corpo. Per una semiotica dell'impronta*, a cura di P. Basso Fossali, Meltemi, Roma, 2004.

Fontanille, J. (2016). *L'Enonciation pratique à l'œuvre dans l'intermédialité et la remédiation*, in Migliore, T. (a cura di), *Rimediazioni. Immagini interattive*, Aracne, Roma, pp. 231-244.

Fontanille, J., & Zilberberg, C. (1998). *Tension et signification*. Mardaga.

Foucault, M. (1969). *L'archéologie du savoir*. Paris: Gallimard; tr. it. *L'archeologia del sapere*.

Foucault, M. (1970). L'ordre du discours: leçon inaugurale au Collège de France prononcée le 2 décembre 1970. Paris: Gallimard; tr. it. Foucault (2001). L'ordine del discorso in Bertani (a cura di), Michel Foucault. Il discorso, la storia, la verità. Einaudi.

Foucault, M. (1976). *Histoire de la sexualité: Vol I—La Volonté de savoir*. Gallimard; tr. It. *Storia della sessualità, vol. I—La volontà di sapere*. Feltrinelli.

Foucault, M. (2010). *Eterotopia*. Mimesis.

Gallese, V., Morelli, U. (2024). *Cosa significa essere umani? Corpo, cervello e relazione per vivere nel presente*. Raffaello Cortina Editore

Goodfellow, I., Bengio, Y., & Courville, A. (2017). *Deep Learning*. MIT Press.

Goodman, A. & Ippolita (2025). *Macchine neurodivergenti. Relazioni postumane e algoritmi queer*. Bologna: Ombre Corte

Greimas, A. J. (1970). *Du sens*. Editions du Seuil; tr. it. *Del Senso I*. Bompiani (1974).

Greimas, A. J., Courtés, J., (1979). *Sémiotique. Dictionnaire raisonné de la théorie du langage*; tr. It. (2007) *Semiotica: Dizionario ragionato della teoria del linguaggio*. B. Mondadori. https://books.google.it/books?id=gXd-FU_xNl8C

Guarino Nicola, 1994. The Ontological Level in Roberto Casati, Barry Smith e Graham White (a cura di), *Philosophy and the Cognitive Sciences*, Hölder-Pichler-Tempsky, Vienna 1994, pp. 443-456.

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14(3), 575–599. <https://doi.org/10.2307/3178066>

Haraway, D. J. (1991). *A Cyborg Manifesto*. Routledge; tr. it. *Manifesto Cyborg. Donne, tecnologie e biopolitiche del corpo*. Feltrinelli (2018).

Harding, S. (1986). *The science question in feminism*. Cornell University Press.

Harris, Z. S. (1951). *Methods in structural linguistics*. Chicago : University of Chicago Press. <http://archive.org/details/methodsinstructu0000harr>

Hayles, N. K. (2006). Unfinished Work: From Cyborg to Cognisphere. *Theory, Culture & Society*, 23(7–8), 159–166. <https://doi.org/10.1177/0263276406069229>

Hayles, N. K. (2019). Can Computers Create Meanings? A Cyber/Bio/Semiotic Perspective. *Critical Inquiry*, 46(1), 32–55. <https://doi.org/10.1086/705303>.

Hegel, G.W.F. (1807), *Phänomenologie des Geistes*; tr. it. De Negri, E. *Fenomenologia dello spirito*, 2 voll., La Nuova Italia editrice (1973).

Hermans, S., Gijzen, M., Mombaers, T., & van Petegem, P. (2022). Gendered patterns in students' motivation profiles regarding iSTEM and STEM test scores: A cluster analysis. *International Journal of STEM Education*, 9(1).

Hjelmslev, L. (1943). *Langue et parole*. In *Essais Linguistiques* (1959). Nordisk Sprog- og Kulturforlag, pp. 69-81.

Hjelmslev, L. (1954). *La stratification du langage*. In *Essais Linguistiques* (1959). Nordisk Sprog- og Kulturforlag.

Hjelmslev, L. (1961), *Prolegomena to a Theory of Language*, tr. it. (a cura di) Lepschy, G.C. (1968), *I fondamenti della teoria del linguaggio*, Einaudi

Houssein, E. H., Saad, M. R., Hashim, F. A., Shaban, H., & Hassaballah, M. (2020). Lévy flight distribution: A new metaheuristic algorithm for solving engineering optimization problems. *Engineering Applications of Artificial Intelligence*, 94, 103731. <https://doi.org/10.1016/j.engappai.2020.103731>

Jakobson, R. (1959). *On linguistic aspects of translation*. In R. A. Brower (Ed.), *On Translation* (pp. 232–239). Cambridge, MA: Harvard University Press.

Jakobson, R. (1963). *Essais de linguistique générale*. Paris: Minuit; Trad. it. *Saggi di linguistica generale*, Feltrinelli, 2002.

Jin, C. and Rinard, M. (2024). Emergent representations of program semantics in language models trained on programs. In Proceedings of the 41st International Conference on Machine Learning (ICML'24), Vol. 235. JMLR.org, Article 891, 22160–22184.

Kaelbling L. P. (2020) *Introduction to Machine Learning*. MIT Opencourseware. <https://youtu.be/kXOsRyIVAdo?si=nPZiDZL3n53-YEuJ>

Kaveh, M., Mesgari, M.S. (2023), Application of Meta-Heuristic Algorithms for Training Neural Networks and Deep Learning Architectures: A Comprehensive Review. *Neural Process Lett* 55, 4519–4622 (2023). <https://doi-org.ezproxy.unibo.it/10.1007/s11063-022-11055-6>

Latour, B. (1998). *Piccola filosofia dell'enunciazione*. Basso P., Corrain L. *Eloqui de senso. Dialoghi semiotici per Paolo Fabbri, Orizzonti, compiti e dialoghi della semiotica*, Costa & Nolan, pp.71-94

Latour, B. (2012). *Enquête sur les modes d'existence*. La Découverte.

Lenci A, Sahlgren M. (2023) From Usage to Meaning: The Foundations of Distributional Semantics. In: Distributional Semantics. *Studies in Natural Language Processing*. Cambridge University Press, p. 3-25.

Leone, M. (2023a). The main tasks of a semiotics of artificial intelligence. *Language and Semiotic Studies*, 9(1), 1–13. <https://doi.org/10.1515/lass-2022-0006>

Leone, M. (2023b). *Semiotica della latenza* in *Il volto latente*. FACETS Digital Press, Open Access, pp. 3-17.

Leone, M. (2025b), *Faut-il être polis avec Chat-GPT?* in Séminaire international de sémiotique à Paris “Intelligence artificielle générative et nouveaux enjeux sémiotiques” organizzato da Dondero, M.G. e Aldama, J.A.

Levitt, H. M. (2025). Attending to investigator reflexivity and epistemic privilege: A worksheet for facilitating individual and research team self-reflection. *Qualitative Psychology*, 12(1), 1–6. <https://doi.org/10.1037/qup0000331>

Levitt, H. M. (2024). A consideration of the influence of researchers’ lived experiences: Reflections on epistemic privilege and social justice in qualitative research. *Qualitative Psychology*, 11(3), 315–327. <https://doi.org/10.1037/qup0000318>

Liu, Y., Ott, M., Goyal, N., ... & Stoyanov, V. (2019). *RoBERTa: A robustly optimized BERT pretraining approach* (arXiv:1907.11692). arXiv. <http://arxiv.org/abs/1907.11692>

Liu, Y., Jun, E., Li, Q., & Heer, J. (2019). Latent Space Cartography: Visual Analysis of Vector Space Embeddings. *Computer Graphics Forum*, 38(3), 67–78. <https://doi.org/10.1111/cgf.1367>

Liu, Y., He, H., Han, T., Zhang, X., Liu, M., Tian, J., Zhang, Y., Wang, J., Gao, X., Zhong, T., Pan, Y., Xu, S., Wu, Z., Liu, Z., Zhang, X., Zhang, S., Hu, X., Zhang, T., Qiang, N., ... Ge, B. (2025). Understanding LLMs: A comprehensive overview from training to inference. *Neurocomputing*, 620, 129190. <https://doi.org/10.1016/j.neucom.2024.129190>

Lobaccaro, L. (2022) *La semiotica cognitiva in Italia. Genesi, diramazioni, evoluzioni*, in: *Cura del senso e critica sociale. Riconoscimento della semiotica italiana*, Roma, Mimesis Edizioni srl, pp. 61 – 94.

Lobaccaro, L. (2024). Acts of enunciation and embodied practices: A cognitive semiotic perspective. *Rivista Italiana Di Filosofia Del Linguaggio*. <https://doi.org/10.4396/SFL20230614>.

Lorusso, M. M. (2025) *Reading distress as oppression and resilience as resistance: a structural ecological analysis of trans and nonbinary people's experiences in Italy*, [Dissertation thesis], Alma Mater Studiorum Università di Bologna. Dottorato di ricerca in Psychology, 38 Ciclo.

Marbut, A. C., Chandler, J. W., & Wheeler, T. J. (2024). *Exploring the Impact of a Transformer’s Latent Space Geometry on Downstream Task Performance*. arXiv:2406.12159.

- Marconi, D. (1987). Two aspects of lexical competence. *Lingua e Stile*, 22(3), 385–395.
- Marconi, D. (1991). Understanding and reference. *Sémiotiques*, 1(1), 9–25.
- Marcus, G. (2024). *No AGI (and no Trustworthy AI) without Neurosymbolic AI*. In *Ontology Summit 2024*.
- Marrone, G. e Migliore, T. a cura di (2021), *La competenza esperta. Tipologie e trasmissione*, Meltemi.
- Martinelli, P. (2025) *Literacy. Semiotica evolutiva dell'apprendimento*. Luca Sossella Editore
- Massumi, B. (2002). *Parables for the virtual: Affect, movement, sensation*. Durham: Duke University Press.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*: August 31, 1955. *AI Mag.*, 27(4), 12–14. <https://doi.org/10.1609/aimag.v27i4.1904>
- Meinke, A., Schoen, B., Scheurer, J., Balesni, M., Shah, R., & Hobbhahn, M. (2025). *Frontier Models are Capable of In-context Scheming* (arXiv:2412.04984). arXiv. <https://doi.org/10.48550/arXiv.2412.04984>
- Mikolov, T., Chen, K., Corrado, G. S., Dean, J. (2013). *Efficient estimation of word representations in vector space*. In: *Proc. ICLR*.
- Mohammadi, A., & Sheikholeslam, F. (2023). Intelligent optimization: Literature review and state-of-the-art algorithms (1965–2022). *Engineering Applications of Artificial Intelligence*, 126, 106959. <https://doi.org/10.1016/j.engappai.2023.106959>
- Monti, N. (2025), *Prompting. Poetiche e politiche dell'intelligenza artificiale*, Tlon.
- Müller, S. (2025). Large Language Models: The best linguistic theory, a wrong linguistic theory, or no theory at all? *Journal of the Linguistic Society of Germany*, 44(1). <https://doi.org/10.18148/ZS/2025-2001>
- Murugesan, S. (2025). The Rise of Agentic AI: Implications, Concerns, and the Path Forward. *IEEE Intelligent Systems*, 40(2), 8–14. <https://doi.org/10.1109/MIS.2025.3544940>

Natale, S. (2021). *Deceitful Media: Artificial Intelligence and Social Life After the Turing Test*. United Kingdom: Oxford University Press, pp. 70-93

Paolucci, C. (2006). Lucien Tesnière autore della logica dei relativi. *E/C*, 2006, 1–16.

Paolucci, C. (2010). *Strutturalismo e interpretazione*. Bompiani.

Paolucci, C. (2017). Sfuggire ai cliché. Gli stereotipi tra enciclopedia, enunciazione e soggettività nel linguaggio. *Reti, saperi, linguaggi*, 2(2017). <https://doi.org/10.12832/88789>

Paolucci, C. (2020). *Persona. Soggettività nel linguaggio e semiotica dell'enunciazione*. Bompiani.

Paolucci, C. (2021a). *Cognitive Semiotics: Integrating Signs, Minds, Meaning and Cognition* (Vol. 24). Springer International Publishing. <https://doi.org/10.1007/978-3-030-42986-7>

Paolucci, C. (2021b). *L'enunciazione come eterogenesi differenziale*. In T. Migliore (a cura di), *Destinatari e destinanti*. Milano: Meltemi.

Paolucci, C. (2021c). L'énonciation comme addition de soustractions : niches sémiotiques, instances énonçantes et sense-making, «*ACTES SÉMIOTIQUES*», 2021, 125, pp. 1 - 13

Paolucci, C. (2023). Pre-Truth: Fake News, Semiological Guerrilla Warfare, and Some Other Media and Communication “Revolutions”. *Media and Communication*, 11(2), 101–108. <https://doi.org/10.17645/mac.v11i2.6628>

Paolucci, C. (2024a). *L'énonciation en tant qu'hétérogénèse différentielle*. In A. Sarti & G. Citti (Eds.), *Dynamiques post-structurelles. Essais sur le devenir des formes*. Paris: Spartacus-idh.

Paolucci, C. (2024b). A semiotic lifeworld. Semiotics and phenomenology: Peirce, Husserl, Heidegger, Deleuze, and Merleau-Ponty. *Semiotica*, 2024(260), 25–43. <https://doi.org/10.1515/sem-2024-0152>

Paolucci, C. (2025a). The myth of meaning: generative AI as language-endowed machines and the machinic essence of the human being. *Semiotica*, 2025(262), 5–23. <https://doi.org/10.1515/sem-2024-0204>

Paolucci, C. (2025b), *Agency, énonciation, signification. Ce que l'IA fait à la sémantique sémiotique*, in Séminaire international de sémiotique à Paris “Intelligence artificielle générative et nouveaux enjeux sémiotiques” organizzato da Dondero, M.G. e Aldama, J.A.

Paolucci, C. (2025c). *Nati cyborg. Cosa l'intelligenza artificiale generativa ci dice dell'essere umano*. Luca Sossella Editore.

Pasquinelli, M (2023). *The Eye of the Master: A Social History of Artificial Intelligence*, London, Verso Books; tr. it. (2025). *Nell'occhio dell'algoritmo. Storia e critica dell'intelligenza artificiale*. Carocci editore (2025).

Pedram, M., Mousavirad, S.J., Schaefer, G. (2022). Training Neural Networks with Lévy Flight Distribution Algorithm. In: Kim, J.H., Deep, K., Geem, Z.W., Sadollah, A., Yadav, A. (eds) *Proceedings of 7th International Conference on Harmony Search, Soft Computing and Applications. Lecture Notes on Data Engineering and Communications Technologies*, vol 140. Springer, Singapore. https://doi-org.ezproxy.unibo.it/10.1007/978-981-19-2948-9_10

Peirce, C. S. (1931–1958). *Collected Papers of Charles Sanders Peirce* (C. Hartshorne, P. Weiss & A. W. Burks, Eds.). Cambridge, MA: Harvard University Press. — Tr. it.: Peirce, C. S. (2005). *Scritti scelti*. Torino: UTET.

Peirce C. S. (1868), Some Consequences of Four Incapacities, in «*Journal of Speculative Philosophy*», 2(1868), pp. 140-157. Ripubblicato in CP 5.264-317 e W2: 211-242. Trad. it., *Scritti scelti*, UTET (2005).

Penfield, C. (2023). Foucault's Virtual Force Ontology. *The Journal of Speculative Philosophy* 37(3), 447-457. <https://muse.jhu.edu/article/904128>

Penge, R., Ubertini, C., Tosco, A., & Piredda, L. (2009). *Abilità narrative e competenze di produzione del testo scritto. Psichiatria dell'infanzia e dell'adolescenza*, 76(3), 315–329.

Pennycook, A., 2018. *Posthumanist applied linguistics*. Routledge.

Perez, C. C., 2020). *Invisibili: Come il nostro mondo ignora le donne in ogni campo. Dati alla mano*. Einaudi.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. In *Proceedings of NAACL-HLT 2018* (pp. 2227–2237).

Petitot, J. (1985). *Morphogenèse du Sens. Pour un schématisme de la structure*. Presses universitaires de France.

Petitot J. (1989a), «Hypothèse localiste, Modèles morphodynamiques et Théories cognitives: Remarques sur une note de 1975», *Semiotica*, 77, 1/3, p. 65-119;

Petitot J. (1989b), «Modèles morphodynamiques pour la Grammaire cognitive et la Sémiotique modale», *RSSI (Canadian Semiotic Association)*, 9, 1-2-3, p. 17-51;

Petitot J. (1990), «Le Physique, le Morphologique, le Symbolique. Remarques sur la Vision», *Revue de Synthèse*, 1-2, p. 139-183;

Petitot J. (1991), «Syntaxe topologique et Grammaire cognitive», *Langages*, 103, p. 97-128.

Petitot, J. (1994). *Attractor syntax: Morphodynamics and cognitive grammar*. In C. Fuchs & B. Victorri (A c. Di), *Linguisticæ Investigationes Supplementa* (Vol. 19, pp. 167–188). John Benjamins Publishing Company. <https://doi.org/10.1075/lis.19.14pet>;

Petitot, J. (1995). *Approche morphodynamique de l'iconicité des stemmas*. In J. Richard-Zappella (a cura di), *Lucien Tesnière aujourd'hui* (pp. 105–112). Peeters.

Petitot, J. (2011). *Cognitive Morphodynamics*. Peter Lang CH. <https://doi.org/10.3726/978-3-0351-0192-8>.

Pezzini, I. (2007). *Il testo galeotto: La lettura come pratica efficace*. Meltemi.

Phang, J., Lampe, M., Ahmad, L., Agarwal, S., Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., & Maes, P. (2025). Investigating Affective Use and Emotional Well-being on ChatGPT. OpenAI.

W.P. Phillips, J. (2006). Agencement/Assemblage. *Theory, Culture & Society*, 23(2-3), 108-109. <https://doi.org/10.1177/026327640602300219>

Piantadosi, S. T., & Hill, F. (2022). *Meaning without reference in large language models* (No. arXiv:2208.02957). arXiv. <http://arxiv.org/abs/2208.02957>

Piantadosi, S. (2024). *Modern language models refute Chomsky's approach to language*. LingBuzz. <https://lingbuzz.net/lingbuzz/007180>

Pirsiavash, H., Vondrick, C., & Torralba, A. (2014). Inferring the Why in Images: Defense Technical Information Center. <https://doi.org/10.21236/ADA612444>

Platone, Filebo, a cura di Migliori, M. Milano 1995.

Politi, F. (2024). Sull'utilità e sul danno della matematica per la comprensione della semiosi macchinica. *E|C — Rivista dell'Associazione Italiana di Studi Semiotici*, 41, 454–463.

Progetto LG-DSA-2018. (2023). *Buone pratiche clinico-assistenziali nella gestione dei DSA: Documento di consenso*. Centro studi Erickson.

Puar, J. K. (2007), *Terrorist Assemblages: Homonationalism in Queer Time*. Duke University Press

Puar, J.K. (2012). "I would rather be a cyborg than a goddess": Becoming-Intersectional in Assemblage Theory. *philoSOPHIA* 2(1), 49-66. <https://dx.doi.org/10.1353/phi.2012.a486621>.

Qi, W., Pan, J., Lyu, H., & Luo, J. (2024). Excitements and concerns in the post-ChatGPT era: Deciphering public perception of AI through social media analysis. *Telematics and Informatics*, 92, 102158.

Qiu, Y., & Jin, Y. (2024). ChatGPT and finetuned BERT: A comparative study for developing intelligent design support systems. *Intelligent Systems with Applications*, 21, 200308. <https://doi.org/10.1016/j.iswa.2023.200308>.

Quillian, R. (1961), A revised design for an understanding machine, in *Mechanical Translation*, Vol.7, no.1, July 1962.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training*. <https://openai.com/index/language-unsupervised/>

Rastier, F. (1999), *Dalla significazione al senso: per una semiotica senza ontologia*, in *Eloquio del senso*, Basso P. e Corrain L. (a cura di), Costa & Nolan, 1999, pp. 213-240.

Roncaglia, G. (2023). *L'architetto e l'oracolo. Forme digitali del sapere da Wikipedia a ChatGPT*. Laterza.

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, Purpose and Teleology. *Philosophy of Science*, 10(1), 18–24.

Rothman, D. (2022), *Transformers for Natural Language Processing: Build, Train, and Fine-tune Deep Neural Network Architectures for NLP with Python, Hugging Face, and OpenAI's GPT-3, ChatGPT, and GPT-4*. Packt Publishing.

Russell, S. J., & Norvig, P. (2022). *Artificial Intelligence: A modern approach* (4th ed.). Pearson.

Salton, Gerard, Wong, Andrew, and Yang, Chung-Shu. 1975. A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620.

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229. <https://doi.org/10.1147/rd.33.0210>

Sanford, C., Hsu, D., & Telgarsky, M. (2024). *Transformers, parallel computation, and logarithmic depth* (No. arXiv:2402.09268). arXiv. <https://doi.org/10.48550/arXiv.2402.09268>

Sarti, A., Citti, G., & Piotrowski, D. (2019). Differential heterogenesis and the emergence of semiotic function. *Semiotica*, 2019(230), 1–34. <https://doi.org/10.1515/sem-2018-0109>

Sarti, A., Citti, G., & Piotrowski, D. (2022). *Differential Heterogenesis: Mutant Forms, Sensitive Bodies*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-97797-9>

Saussure, F. de. (1922). *Cours de linguistique générale*. Paris: Payot. — Trad. it.: Saussure, F. de. (2003). *Corso di linguistica generale* (T. De Mauro, a cura di). Roma-Bari: Laterza.

Shavit, Y., et al. (2023). *Practices for Governing Agentic AI Systems*. OpenAi.

Schoen, B., Nitishinskaya, E., Balesni, M., Højmark, A., Hofstätter, F., Scheurer, J., Meinke, A., Wolfe, J., Weij, T. van der, Lloyd, A., Goldowsky-Dill, N., Fan, A., Matveiakina, A., Shah, R., Williams, M., Glaese, A., Barak, B., Zaremba, W., & Hobbhahn, M. (2025). *Stress Testing Deliberative Alignment for Anti-Scheming Training* (No. arXiv:2509.15541). arXiv. <https://doi.org/10.48550/arXiv.2509.15541>

Sforzini, A. (2014). *Michel Foucault entre histoire et fiction*. In M. Panter, P. Mounier, M. Martinat, & M. Devigne (eds.), *Imagination et histoire: enjeux contemporaine*. Presses universitaires de Rennes. <https://doi.org/10.4000/books.pur.49586>

Simondon, G. (1964), *L'Individuation à la Lumière des Notions de Forme et d'Information*, Presses Universitaires de France; tr. it. *L'individuazione alla luce delle nozioni di forma e d'informazione*, Mimesis (2011).

Somaini, A. (2023). *Algorithmic images: Artificial intelligence and visual culture*. Grey Room, 93, 74–115. https://doi.org/10.1162/grey_a_00383

Spivak, G. C. (1988). *Can the Subaltern Speak? Marxism and the Interpretation of Culture*, 271–313.

Sowa, J. (2024). *Without Ontology, LLMs are clueless*. In Ontology Summit 2024.

Stano, S. (2023) *Critique of Pure Nature*. Cham: Springer. ISBN (Hardcover): 978-3-031-45074-7, ISBN (Softcover): 978-3-031-45077-8, e-ISBN: 978-3-031-45075-4 [EN].

Suoizzi, A., Capone, L., Lebani, G. E., & Lenci, A. (2025). *BAMBI: Developing Baby Language Models for Italian* (No. arXiv:2503.09481). arXiv. <https://doi.org/10.48550/arXiv.2503.09481>

Sütiste, E. (2021). *Intersemiotic translation*, in Gambier, Y., & Van Doorslaer, L. (A c. Di). (2021). *Handbook of Translation Studies* (Vol. 5). John Benjamins Publishing Company, pp. 127–133. <https://doi.org/10.1075/hts.5>.

Terranova, T. (2022). *After the Internet: Digital networks between capital and the common*. Semiotext(e). — Trad. it.: Terranova, T. (2024). *Dopo Internet. Le reti digitali tra capitale e comune*. Nero.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris: Klincksieck. (Rif. a stemmi/ordine strutturale).

Thom, R. (1977), *Stabilité structurelle et morphogénèse. Essai d'une théorie générale des modèles*, InterEditions, Paris. Tr. It. *Stabilità strutturale e morfogenesi. Saggio di una teoria generale dei modelli*, Ghibli (2023).

Timeto, F. (2020). *Bestiario Haraway Per un femminismo multispecie*. Mimesis

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind, New Series*, 59(236), 433–460

Valle, A. (2003). Le due facce del senso. Note su espressione e contenuto. *Semiotiche*, 1, 2003, pp. 13–43.

Valle, A. (2025). From grammar to text: A semiotic perspective on a paradigm shift in computation and its usages. *Semiotica*. <https://doi.org/10.1515/sem-2024-0187>

Varela, F. J., & Thompson, E. (2000). *Neural Synchrony and the Unity of Mind: A Neurophenomenological Perspective*.

Vaswani, A., et al. (2017). *Attention Is All You Need*. NeurIPS 2017. <http://arxiv.org/abs/1706.03762>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (No. arXiv:1706.03762). arXiv. <http://arxiv.org/abs/1706.03762>

Vetere, G. (2025). *Intelligenze aliene. Linguaggio e vita degli automi*. Luca Sossella Editore.

Vig, J. (2019). A Multiscale Visualization of Attention in the Transformer Model. In M. R. Costa-jussà & E. Alfonseca (A c. Di), *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 37–42). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-3007>.

Violi, P. (1997). *Significato ed esperienza*, Milano, Bompiani.

Violi, P. (2003). Significati lessicali e pratiche comunicative. Una prospettiva semiotica. *Rivista di linguistica*, 15(2), 321-342.

Violi, P. (2015). Global and local: Encyclopedic meaning revisited. *Semiotica*, 2015(206). <https://doi.org/10.1515/sem-2015-0023>.

Violi, P. (2017). *Encyclopedia: Criticality and actuality* in Sara G. Beardsworth, Randall E. Auxier (2017) *The Philosophy of Umberto Eco. Library of Living Philosophers*. Open Court; tr. it. Violi (2021), *L'Enciclopedia: criticità e attualità*, in *La filosofia di Umberto Eco*. La nave di Teseo.

Voto, C. (2021). Opacizzare il volto artificiale attraverso le arti digitali: Errori, deformità, materia, intersoggettività. *Lexia, Volti artificiali* (37–38), 285–302.

Voto, C. (2023). *Verso una semiotica della spazialità latente*, in Leone, M. (2023), *Il volto latente*, FACETS Digital Press, Open Access.

Waddington, C. H. (1957). *The Strategy of the Genes*. Routledge. <https://doi.org/10.4324/978131576547>

Weizenbaum, J. (1961) How to Make a Computer Appear Intelligent. *Datamation* 7, pp. 24–26.

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>

Whalen, T. (2000). *Data Navigation, Architectures of Knowledge*. Paper presented at the Banff Summit on Living Architectures: Designing for Immersion and Interaction, Banff New Media Institute, 23 Sept

Zanelli, S. (2025). *Deleuze con Peirce. Un equivoco impossibile*. Orthotes

Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., & Fidler, S. (2015). Aligning Books and Movies: Towards Story-Like Visual Explanations by Watching Movies and Reading Books. 2015 *IEEE International Conference on Computer Vision (ICCV)*, 19–27. <https://doi.org/10.1109/ICCV.2015.11>

Sitografia

Commissione Europea – DigCompEdu. European Framework for the Digital Competence of Educators (2017). https://joint-research-centre.ec.europa.eu/digcompedu_en

Data Europa EU (doi). Orientamenti etici per gli educatori... (2022).
<https://data.europa.eu/doi/10.2766/438>

Eredità del “Trattato di semiotica generale” di Umberto Eco:
<https://www.youtube.com/watch?v=gSEEdM7k9A>

[FMSH] Fondation Maison des sciences de l’homme. Séminaire international de sémiotique – Paris 3. <https://www.fmsh.fr/projets/seminaire-international-de-semiotique-paris-3>

Garante per la protezione dei dati personali (30 gennaio 2025). Provvedimento.
<https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10098477>

Incontri di Quinto Alto (Regista). (2023, marzo 14). A Sarti «Differential heterogenesis. Mutant forms, sensitive bodies.» [Film]. <https://www.youtube.com/watch?v=ibUSkIBskjE>

Italia Domani (s.d.). Assegnazione di borse di dottorato in collaborazione con le imprese e promozione dell’assunzione dei ricercatori dalle imprese.
<https://www.italiadomani.gov.it/content/sogei-ng/it/it/Interventi/investimenti/Assegnazione-di-borse-di-dottorato-in-collaborazione-con-le-imprese-e-promozione-dellassunzione-dei-ricercatori-dalle-imprese.html>

Kaggle – Amazon Fine Food Reviews. <https://www.kaggle.com/datasets/snap/amazon-fine-food-reviews>

OpenAI Tokenizer: <https://platform.openai.com/tokenizer>

PARTS: Persona, Aim, Recipients, Theme, and Structure.
<https://edu.exceedlms.com/student/path/1721586/activity/3310073#/page/65de4aee478a58477d288a2f>
https://www.repubblica.it/tecnologia/2025/08/11/news/rivolta_degli_utenti_contro_chatgpt_5_o_penai_riattiva_i_vecchi_modelli-424783672/

VADER Sentiment – Documentation. <https://vadersentiment.readthedocs.io/en/latest/>

Weng, L. *Attention? Attention!* <https://lilianweng.github.io/posts/2018-06-24-attention/>

Abstract

Italiano

La tesi propone una teoria della semiosi algoritmica applicata ai Large Language Models (LLM) che: i) assume il prompt come topologia dinamica capace di attivare il virtuale intensivo enciclopedico, nel suo carattere globale altrimenti irrepresentabile, consentendone un ritaglio locale; ii) legge la capacità di generazione di nuovi enunciati come effetto di un assemblaggio tra istanze eterogenee. Il lavoro intreccia semiotica interpretativa e cognitiva con elementi di epistemologia, filosofia del linguaggio, matematiche, scienze computazionali e studi di genere al fine di chiarire statuti ontologici ed epistemologici dei testi generati, nonché i limiti e le potenzialità delle macchine dotate di linguaggio. A un inquadramento generale utile per definire in che termini si possa parlare di senso, significazione e competenze semiotiche a proposito dei modelli generativi, seguono due sguardi prospettici sul loro funzionamento: da un lato, i processi di significazione dell'architettura transformer (*IA in sé*); dall'altro l'ibridazione cyborg tra umano e macchina (*IA per sé*). L'enunciazione impersonale ed evenemenziale e l'eterogenesi differenziale vengono poste a fondamento teorico come dispositivi euristici capaci di tenere insieme questi due sguardi combinati. Viene, in conclusione, dedicato un capitolo alla filiera formativa STEM da una prospettiva di genere e presentato il percorso di ricerca in impresa condotto presso ART-ER società consortile della regione Emilia-Romagna.

Inglese

The dissertation proposes a theory of algorithmic semiosis applied to Large Language Models (LLMs) that: i) conceptualizes the prompt as a dynamic topology capable of activating the encyclopedic intensive virtual, in its otherwise globally unrepresentable character, thereby enabling a local extraction of knowledge; ii) interprets the capacity to generate new enunciates as the effect of an assemblage among heterogeneous instances. The work interweaves interpretive and cognitive semiotics with elements of epistemology, philosophy of language, mathematics, computational sciences, and gender studies in order to clarify the ontological and epistemological status of generated texts, as well as the limits and potentials of language-endowed machines. Following a general theoretical framing aimed at defining the terms in which sense, signification, and semiotic competences may be discussed with regard to generative models, the dissertation offers two perspectives on their functioning: on the one hand, the processes of signification within the transformer architecture (*AI in itself*); on the other, the cyborg-hybridization between human and machine (*AI for itself*). Impersonal and evenemential enunciation, together with differential

heterogenesis, are established as the theoretical foundations, that is heuristic devices capable of holding together these two combined perspectives. The dissertation concludes with a chapter devoted to the STEM educational pipeline from a gender perspective and presents the results of an industrial research placement conducted at ART-ER, the consortium company of the Emilia-Romagna region.