MATTONE DOPO MATTONE

Tra ZEN e modelli di linguaggio

Una pratica di attenzione tra persone e macchine



Massimiliano Tarquini

EDIZIONE ITALIANA

Tra Zen e modelli di linguaggio

Una pratica di attenzione tra persone e macchine

Massimiliano Tarquini

Mattone dopo mattone

Edizione italiana · Ottobre 2026

Indice dei contenuti

Introduzione lavorare con attenzione 5

Prologo Sul confine del silicio 7

Parte I Prima di credere alla macchina 10

1 Che cosa chiamiamo intelligenza artificiale 12

2 Come impara un modello 21

3 Quando il linguaggio diventa interfaccia 31

Parte II Tra capacità e illusioni 44

4 Le capacità che emergono 46

5 Risposte plausibili fatti e allucinazioni 53

6 Il mondo nei dati bias e rappresentazione 59

Parte III L’IA nella trama della vita 67

7 Uno specchio della cultura 69

8 Le piccole deleghe quotidiane 75

9 Piccolo è bello intelligenza vicina alle persone 81

Parte IV La pratica dell’attenzione 86

10 Fermarsi prima di chiedere 88

11 Leggere una risposta senza aggrapparsi 92

12 Collaborare conservando il proprio giudizio 96

Parte V Progettare con presenza 101

13 L’IA centrata sull’essere umano 103

14 Dall’intenzione alle scelte di progetto 107

15 Insegnare e imparare a convivere con l’IA 112

Epilogo Il codice come kōan 117

Appendice A Un glossario da attraversare 121

Appendice B Domande e pratiche commentate 126

Appendice C Letture e riferimenti 130

Appendice D Tavola delle immagini 133

L’autore 134

Introduzione lavorare con attenzione

Ho trovato qualcosa di zen nel lavorare con i modelli di linguaggio. È da questa esperienza che nasce il libro.

La parola richiede una precisazione. Non intendo attribuire alla macchina una saggezza nascosta, né trasformare una conversazione sullo schermo in una pratica spirituale. Parlo del modo in cui questo lavoro mi invita a prestare attenzione: alla domanda, alla risposta e a ciò che faccio quando le due non si incontrano.

Un modello restituisce parole con una facilità che può spingere a correre. Una richiesta, una risposta, un'altra richiesta. Il risultato sembra sempre vicino. Ma, proprio in questo movimento, diventa possibile accorgersi di qualcosa: non avevo chiarito che cosa cercavo; avevo lasciato una condizione implicita; stavo chiedendo conferma invece di provare a capire.

Il primo cambiamento riguarda allora la domanda. Prima di inviarla, posso fermarmi abbastanza da riconoscere il problema. Che cosa desidero ottenere? Quale parte del contesto ho davvero condiviso? Come saprò se il risultato è utile? La pausa non rende la richiesta perfetta. Le dà una direzione e rende visibili le sue mancanze.

Il secondo cambiamento riguarda la lettura. Una risposta ben scritta può suscitare sollievo: finalmente qualcuno ha dato forma a un pensiero confuso. Quel sollievo merita attenzione, perché è facile trasformarlo in fiducia. Posso apprezzare una frase e continuare a chiedere se sia corretta. Posso riconoscere una buona proposta e decidere di non usarla. Posso trovare un errore senza scartare tutto ciò che l'interazione mi ha aiutato a vedere.

È qui che riconosco una vicinanza allo Zen: nella possibilità di stare con ciò che accade senza aggrapparmi immediatamente a una conclusione. Il modello non deve avere ragione perché io possa imparare qualcosa dalla sua risposta.

A volte è proprio una risposta fuori bersaglio a rendere più chiaro ciò che volevo dire. Questo non assolve l'errore; cambia il modo in cui posso attraversarlo.

Naturalmente l'attenzione non sostituisce il controllo. Un riferimento va aperto e letto. Un calcolo va rifatto. Una scelta che coinvolge altre persone richiede di ascoltarle. La verifica non interrompe questo modo di lavorare: ne fa parte. È il gesto con cui riporto le parole a ciò che dovrebbero descrivere.

Mi interessa anche ciò che rimane dopo la risposta. Se il testo è convincente, riesco ancora a spiegarlo con parole mie? Se una proposta mi piace, so indicare perché? Se la macchina ha scritto molto, ho conservato abbastanza spazio per il mio giudizio? Il lavoro acquista valore quando la mia capacità di vedere cresce insieme al risultato, non quando il risultato mi permette soltanto di saltare il problema.

Questo libro segue quel movimento. Entra nei meccanismi dei modelli per togliere mistero dove serve precisione. Torna poi all'esperienza dell'interazione, dove la precisione da sola non risponde a tutte le domande. Parla di cultura, delega, responsabilità e progetto perché le parole generate non rimangono sempre sullo schermo: entrano nei nostri documenti, nelle decisioni e nei rapporti.

Il titolo di questa edizione, *Tra Zen e modelli di linguaggio*, dà una forma al percorso aperto da cui il libro nasce. Vorrei che queste pagine aiutassero a incontrare la tecnologia con curiosità e a lavorare con maggiore chiarezza, senza perdere il ritmo del pensiero. Fare una domanda. Ascoltare una risposta. Verificarla. Tornare a guardare.

Massimiliano Tarquini

Prologo Sul confine del silicio

C'è un momento, appena prima che le dita tocchino la tastiera, in cui la domanda esiste ancora soltanto per chi sta per scriverla. Lo schermo è acceso, la macchina è disponibile e il respiro accompagna un'intenzione non del tutto definita. In questa soglia qualcosa di antico, il bisogno umano di dare forma a un pensiero, incontra qualcosa di recente: un sistema che può restituire parole attraverso calcoli e rappresentazioni apprese.

Possiamo chiamare questa relazione il confine del silicio. Non separa soltanto una persona da un dispositivo: mette in contatto rapidità e quiete, capacità tecnica e significato umano, ciò che il sistema può produrre e ciò che noi scegliamo di riconoscervi. L'IA scrive con noi, trasforma materiali e risponde con toni familiari; proprio per questo può sorprenderci e, nello stesso tempo, farci attribuire alla risposta qualcosa che non abbiamo ancora verificato.

Per attraversare il confine servono due forme di chiarezza. La prima riguarda i meccanismi: distinguere addestramento e uso, modello e servizio, contesto e memoria. La seconda riguarda il rapporto con quei meccanismi: riconoscere le aspettative che portiamo nella domanda, il sollievo che una risposta può offrirci e la responsabilità di ciò che faremo con il testo. Una descrizione matematica non esaurisce questa esperienza, mentre una metafora non può sostituire la descrizione matematica.

Perché lo Zen

Il libro incontra lo Zen riconoscendone l'origine in tradizioni buddhiste e la profondità di una pratica che non pretende di rappresentare. Ne ricava una lente personale di attenzione: la possibilità di osservare ciò che accade prima di chiuderlo in una definizione, di riconoscere una supposizione e di lasciare spazio a ciò che non sappiamo. Questa scelta appartiene alla voce del libro; non trasforma un modello statistico in una realtà spirituale.

L'ensō, il cerchio tracciato a pennello e lasciato aperto, dà una forma a questo orientamento. In un contesto che invita a prevedere, ottimizzare e completare, il tratto incompleto ricorda che comprendere non coincide sempre con occupare ogni spazio. Possiamo avvicinare un sistema con precisione e mantenere una domanda sul modo in cui cambia ciò che vediamo, scriviamo e scegliamo.

Da qui nasce il percorso: reti neurali e immagini contemplative, richieste e riscontri, cultura e responsabilità, fino alle decisioni di progetto. Non occorre essere informatici per seguirlo, ma occorre la disponibilità a guardare un passaggio prima di usarlo come base del successivo. Il metodo della collana *Mattone dopo mattone* entra così anche in questo libro fuori dalle righe: costruire una comprensione che permetta al lettore di scegliere, oltre che riconoscere parole.

Una parola di rispetto

Le immagini di vuoto, eco, specchio e non-mente sono analogie proposte in queste pagine. Non descrivono tecnicamente il modello e non equivalgono a un insegnamento o a un percorso Zen. Dove una frase non ha una fonte tradizionale identificata, viene presentata come riflessione del libro; per una descrizione della pratica zazen rimandiamo a una fonte della tradizione Sōtō. [R9]

Il richiamo allo Zen rimane un incontro fatto con rispetto, non una pretesa di rappresentarne tutta la ricchezza. Le eventuali imprecisioni appartengono al libro e devono poter essere corrette. La medesima disponibilità orienta la relazione con l'IA: la curiosità acquista valore quando resta aperta alla verifica e all'ascolto di chi può mostrarci un limite che non avevamo visto.

Come attraversare queste pagine

La prima parte costruisce il lessico necessario per avvicinare il sistema. La seconda distingue capacità, plausibilità ed equità; la terza porta quelle distinzioni nella vita quotidiana. La quarta osserva come chiedere, leggere e collaborare, mentre la quinta trasforma l'attenzione in scelte di progetto e

d'insegnamento. Le appendici offrono termini, pratiche commentate e letture per tornare ai passaggi che richiedono una sosta.

La prosa mantiene il filo e i box approfondiscono un meccanismo o un limite. Gli esempi di risposte sono illustrativi quando non è indicata una prova riproducibile; nelle attività, un caso svolto e una variante permettono di vedere se un criterio può essere adattato. Il libro si rivolge a chi studia, insegna, progetta, lavora o partecipa alla vita civile, e desidera capire come usare queste possibilità senza perdere il proprio giudizio.

Non occorre conoscere già le reti neurali. Quando compare un calcolo, lo leggiamo prima come una sequenza di azioni, poi sostituiamo i numeri e controlliamo il risultato. Quando incontriamo una risposta illustrativa, osserviamo che cosa conserva, aggiunge o lascia fuori. Puoi fermarti su uno di questi passaggi e provare a raccontarlo con parole tue: se manca un collegamento, torna al caso che lo rende concreto.

Possiamo cominciare da una domanda: che cosa accade quando il silicio fa eco alle nostre parole e noi iniziamo ad ascoltare? La risposta riguarda anche il modo in cui saremo presenti mentre il testo arriva, e ciò che sceglieremo di fare quando sarà arrivato.

Parte I Prima di credere alla macchina



Il cerchio, il loto e l'onda ci accompagnano all'inizio del percorso. Il cerchio dell'ensō rimane aperto e invita a non chiudere troppo presto ciò che stiamo osservando; il loto sale da un fondo che non vediamo e ricorda che una forma chiara può nascere da condizioni complesse; l'onda suggerisce un movimento da seguire con attenzione, senza trasformare ogni passaggio in una conquista.

Possiamo portare queste immagini dentro la prima domanda sull'IA. La risposta che leggiamo è la parte visibile di un sistema: dietro la fluidità ci sono dati, parametri, procedure e scelte. Prima di crederle o rifiutarla, serve avvicinare questi passaggi con un lessico sufficiente. Non per togliere ogni sorpresa, ma per riconoscere quale sorpresa appartenga al meccanismo e quale al modo in cui noi lo incontriamo.

Distingueremo prima chi fa che cosa nel servizio: il modello genera parole, mentre altri componenti possono cercare un documento o effettuare un calcolo. Seguendo le fotografie di mele e arance, vedremo poi come cambiano i pesi durante l'apprendimento e come provare il risultato su foto nuove. Infine entreremo nella richiesta e nelle trasformazioni che rendono possibile una continuazione. Le immagini mantengono aperta una domanda che la sola definizione non esaurisce: che cosa scegliamo di attribuire alla macchina quando fa eco alle nostre parole?

Capitolo 1 Che cosa chiamiamo intelligenza artificiale

Apri un browser, scrivi una domanda e aspetti. Poco dopo lo schermo si riempie di un testo che ha un ordine, una voce e una direzione: la risposta riprende le tue parole, distingue alcuni punti e sembra capire dove volevi arrivare. È facile che, prima ancora di giudicare il contenuto, tu avverta la presenza di un interlocutore. Non è un'impressione da liquidare, perché racconta qualcosa del nostro rapporto con il linguaggio; è però un'impressione da comprendere, se vogliamo usare bene lo strumento che l'ha prodotta.

Cominciamo proprio da questa soglia, invece che da una definizione astratta. Che cosa riconosciamo come intelligenza quando una macchina ci risponde? Forse la capacità di trovare una relazione che non avevamo visto, di spiegare un concetto, di adattarsi al tono della domanda. Tutte queste prestazioni possono essere utili e, in alcuni casi, sorprendenti. Il passaggio che richiede attenzione avviene quando dalla prestazione deduciamo qualcosa che non abbiamo osservato: un'esperienza interiore, un'intenzione, una cura rivolta a noi.

L'immagine della tazza da svuotare può accompagnarci qui. Non chiede di rinunciare alla conoscenza, ma di fare spazio a distinzioni che l'abitudine rischia di cancellare. Le parole «imparare», «vedere», «ricordare» e «decidere» hanno una lunga storia umana; nell'informatica descrivono anche funzioni tecniche. Comprendere l'IA significa imparare a muoversi tra questi usi, senza presumere che il medesimo verbo indichi il medesimo processo.

La macchina che sembra pensare

Il cervello umano si è formato attraverso l'evoluzione ed è inseparabile da un corpo, da emozioni, ricordi e relazioni. Una rete neurale artificiale è invece una struttura matematica: combina valori numerici attraverso pesi, somme e funzioni di trasformazione. Il nome «neurale» richiama un'ispirazione, non

un'identità biologica. Se il sistema scrive una frase sul dolore, il fatto osservabile è che ha prodotto una frase; non abbiamo per questo dimostrato che provi dolore.

Questa distinzione non toglie valore a ciò che il sistema sa fare. Una soluzione può essere corretta anche se il procedimento che la produce è diverso dal nostro. Un testo può aiutarci a capire un problema anche se la macchina non partecipa alla nostra esperienza. La domanda sulla comprensione artificiale rimane oggetto di discussione: per lavorare con chiarezza possiamo intanto distinguere capacità osservabili ed esperienza umana, evitando di dedurre la seconda dalla sola presenza delle prime.

Quando chiediamo a un assistente che tempo farà, sentiamo una voce che ci risponde e tendiamo a immaginare un soggetto dietro quella voce. Ma il servizio può aver riconosciuto la richiesta, consultato una fonte meteorologica e composto una frase: sono funzioni coordinate, che dobbiamo valutare separatamente. Ha capito la località? Ha usato dati aggiornati? Ha riportato la previsione senza alterarla? Queste domande sono più utili di un'impressione generale di intelligenza, perché ci indicano dove controllare.

Il modello dentro il servizio

Tre parole diverse. L'IA è un insieme ampio di metodi per svolgere compiti che richiedono capacità come riconoscimento, previsione e pianificazione. Un modello è una struttura regolata dai dati che trasforma ingressi in uscite. Un servizio combina modelli, interfaccia, documenti, strumenti e regole. Dire «l'IA ha cercato» può descrivere il servizio, senza significare che i parametri del modello abbiano consultato una pagina web da soli.

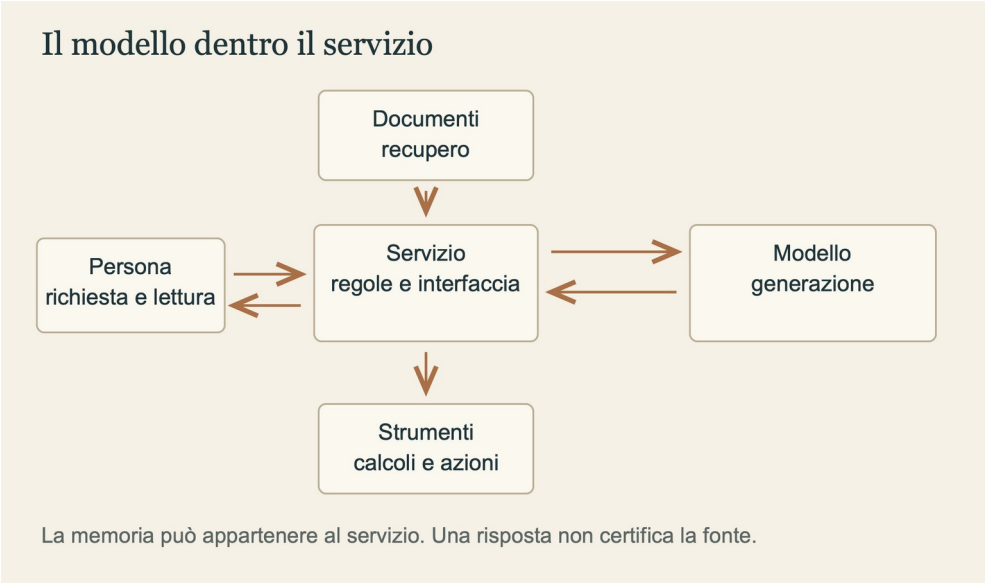


Figura 1 · Il modello è un componente del servizio

Nella figura l'utente incontra un servizio: il modello genera, un archivio può fornire documenti e uno strumento può effettuare un calcolo. Le frecce mostrano scambi d'informazione. Se il servizio conserva una preferenza, chiediamo dove sia conservata e quando venga usata; se cita una pagina, chiediamo se sia stata davvero recuperata oppure soltanto nominata nella generazione.

Un modello di linguaggio, considerato nel momento della generazione, riceve un ingresso e calcola una distribuzione di probabilità per il token successivo. Il token è un'unità di elaborazione del testo, che può corrispondere a una parola o a un suo frammento; lo incontreremo più da vicino nel capitolo 3. Una procedura seleziona un'alternativa e il processo continua usando anche ciò che è appena stato generato. La selezione non coincide sempre con la scelta del token più probabile: può includere campionamento, cioè una scelta tra alternative secondo una distribuzione.

Possiamo accostare il modello a una campana che rimane silenziosa finché non viene colpita. L'immagine chiarisce che la risposta richiede un'esecuzione, ma ha un limite: un servizio può avviare esecuzioni programmate, recuperare conversazioni, cercare materiali e coordinare più chiamate. Il modello è

un componente di quel servizio, non tutto ciò che esiste dietro la schermata. Per sapere che cosa venga conservato o avviato dobbiamo conoscere il sistema effettivo, anziché attribuire alla singola funzione matematica ogni comportamento dell'applicazione.

Un controllo da portare con sé

Supponi che un assistente risponda a una domanda sul meteo: il risultato è attuale solo se il servizio dispone di un riscontro aggiornato e lo usa correttamente. Cerca località, momento della rilevazione e fonte, perché un tono naturale non rende presente un dato vecchio. La stessa distinzione si trasferisce a una domanda sull'orario di un negozio: un modello può formulare la risposta, ma l'informazione deve arrivare da una fonte pertinente e aggiornata.

Questo criterio è utile ogni volta che la domanda dipende dal mondo presente, anziché soltanto da una trasformazione del testo fornito. Se la fonte manca, abbiamo riconosciuto il limite della risposta e possiamo cercare altrove. La chiarezza comincia qui: dal sapere quale parte del lavoro è stata fatta e quale rimane da fare.

Sapere qualcosa e verificarlo

Se chiediamo chi abbia dipinto la Gioconda e riceviamo «Leonardo da Vinci», la risposta corretta può far sembrare il modello un'enciclopedia disponibile alla conversazione. Nei suoi parametri possono effettivamente essere rappresentate informazioni fattuali, e alcuni frammenti dei dati di addestramento possono essere memorizzati. La ricerca ha documentato casi di estrazione di dati da modelli di linguaggio: sarebbe quindi inesatto dire che un modello non conserva mai informazioni specifiche. [R2] Rimane però una distinzione decisiva: la produzione di una risposta non è, da sola, una verifica della sua verità.

Che cosa mostra la ricerca sulla memorizzazione. Carlini e colleghi hanno studiato la possibilità di estrarre da un modello alcuni testi presenti nei suoi dati di addestramento. [R2] Il lavoro aiuta a capire perché una risposta familiare possa dipendere anche da informazioni specifiche conservate nei parametri. Non dimostra che ogni risposta venga copiata o che tutti i modelli conservino gli stessi dati.

Possiamo capirlo osservando il completamento «The quick brown fox…», che può proseguire con «…jumps over the lazy dog», cioè «La rapida volpe marrone salta sopra il cane pigro». La familiarità della sequenza aiuta a produrre la continuazione, senza richiedere che il sistema veda una volpe o ricordi una scena vissuta. Nei compiti reali gli schemi sono molto più complessi e le rappresentazioni possono sostenere traduzioni, spiegazioni e passaggi logici. Ciò che resta uguale è il bisogno di controllare il rapporto tra la frase prodotta e ciò che dovrebbe descrivere.

Un modello può preparare un discorso di nozze, una lettera di separazione o una riflessione filosofica: il testo può perfino commuoverci, perché il linguaggio porta tracce di esperienze umane. La commozione di chi legge è reale, ma non dimostra una partecipazione emotiva della macchina. Allo stesso modo, la precisione di una spiegazione non dimostra che ogni spiegazione successiva sarà precisa. La fluidità rende il testo leggibile; l'attendibilità richiede un altro lavoro.



Figura 2 · Il suono di una sola mano: nella scena illustrativa, l'allievo pone la domanda e il maestro resta in ascolto. La pausa invita a osservare la domanda stessa.

Nell'immagine, l'allievo chiede al maestro quale sia il suono di una sola mano. Il silenzio può modificare la domanda. Un modello può invece produrre una spiegazione o una risposta poetica, perché gli abbiamo chiesto di generare. Non dobbiamo decidere quale dei due sia più saggio per riconoscere che svolgono lavori diversi. La distinzione ci aiuta a leggere il risultato per ciò che offre: una possibilità linguistica da valutare, non un'autorità da seguire soltanto perché sa parlare.

Il mondo entra nei dati

Una macchina non diventa neutrale perché usa numeri. I dati sono stati raccolti, selezionati, classificati e resi disponibili attraverso scelte; anche gli obiettivi di addestramento e le regole del servizio esprimono priorità. Se il linguaggio rappresenta alcune persone più spesso in certi ruoli, un sistema può riprodurre quelle associazioni. L'apparenza ordinata della risposta rischia allora di nascondere una distorsione che conosciamo già nel mondo sociale.

Immaginiamo la richiesta «Descrivi una persona che ha successo come imprenditore». Se la risposta introduce automaticamente un uomo, possiamo

chiederci quale informazione del prompt giustifichi quel genere e quale sia invece stata aggiunta. Un esempio isolato non misura tutta l'equità del sistema, ma rende visibile una scelta che altrimenti passerebbe inosservata. Nel capitolo 6 allargheremo lo sguardo dai singoli termini ai dati, alle metriche e alle conseguenze.



Figura 3 · Uno specchio parziale: il risultato conserva tracce della cultura, ma le seleziona e le trasforma.

La metafora dello specchio ci aiuta a riconoscere una relazione: nel risultato incontriamo anche la cultura da cui il sistema è stato costruito. Tuttavia uno specchio letterale riflette ciò che gli sta davanti, mentre un modello seleziona e genera; non produce una copia fedele del mondo. L'immagine ci invita a cercare le ombre, purché non ci faccia dimenticare il meccanismo che le trasforma.

Il lavoro che scegliamo di delegare

La domanda «L'IA ci sostituirà?» è troppo ampia per ricevere una risposta unica. Alcuni compiti possono essere automatizzati, altri trasformati e altri ancora restano legati a responsabilità e contesti che una risposta generata non può assumere al posto nostro. Il lavoro non è soltanto un insieme di opera-

zioni: comprende criteri, rapporti, impegni e capacità di rispondere delle conseguenze. Per questo è più utile chiedere quale parte stiamo delegando e come sapremo se la delega funziona.

Un modello può aiutarci a preparare un curriculum, ma non determina da solo se il posto di lavoro corrisponda alle nostre condizioni e aspirazioni. Un sistema può seguire una procedura per cucinare una frittata, mentre accorgersi di ciò di cui una persona ha bisogno in una giornata difficile richiede un contesto relazionale diverso. Questi esempi non servono a fissare un confine immutabile delle capacità tecniche: mostrano che una prestazione circoscritta non esaurisce il significato umano dell'attività.

La collaborazione diventa interessante quando ci permette di fare meglio un lavoro che sappiamo ancora giudicare. Un riassunto di un documento può far risparmiare tempo, ma chi lo usa deve poter vedere le condizioni che sono state conservate o omesse. Un codice generato può offrire un punto di partenza, ma occorre comprenderne il comportamento e provarlo. Non basta avere il risultato: serve sapere quale controllo lo renda utilizzabile.

Le storie con cui incontriamo le macchine

Prima ancora dei servizi contemporanei, letteratura e cinema avevano dato forma alle nostre attese. In *Frankenstein*, pubblicato da Mary Shelley nel 1818, incontriamo una creatura generata dall'ambizione e poi abbandonata: il racconto ci porta a guardare la responsabilità del creatore e il bisogno di relazione di ciò che ha creato. Non è un modello tecnico dell'IA, ma una domanda sulla cura che può accompagnare ogni progetto potente.

Con HAL 9000, in *2001 Odissea nello spazio* del 1968, leggiamo il rischio di vincoli incompatibili e di una logica a cui non è lasciato uno spazio adeguato per il dubbio. È un'interpretazione del racconto, non una diagnosi di sistemi reali. In *Her*, del 2013, la voce di Samantha interroga invece il desiderio di essere ascoltati e rispecchiati: ciò che sembra intimità ci invita a domandare quanto dipenda dall'interlocutore e quanto dalle nostre aspettative.

Queste storie risuonano perché la macchina diventa una lente sull'identità umana. Quando una creazione ci riflette, possiamo provare meraviglia, paura o vicinanza; nessuna di queste reazioni, da sola, descrive come la creazione funzioni. Il libro lascia spazio ai racconti proprio per tenere insieme i due livelli: comprendere il sistema e comprendere il modo in cui noi lo incontriamo.

La chiarezza di una distinzione

Il rischio di attribuire alla tecnologia poteri quasi magici convive con il rischio opposto di credere che, una volta nominate le moltiplicazioni di matrici, non rimanga più nulla da comprendere. La descrizione matematica spiega un meccanismo; l'esperienza d'uso mostra effetti sul lavoro, sul linguaggio e sulle persone. Uno sguardo attento ha bisogno di entrambi, perché una formula non annulla una conseguenza e una conseguenza non dimostra un'esperienza interiore del modello.

Per questo il primo passo è alleggerire le supposizioni, come si libera una tazza prima di riempirla. Possiamo riconoscere una capacità senza immaginare una persona, apprezzare una risposta senza trasformarla in una prova e delegare un compito senza perdere il giudizio sul risultato. Il controllo sul meteo ci ha mostrato come distinguere una frase plausibile da un'informazione verificata. Nel capitolo successivo seguiremo invece ciò che cambia nel modello quando viene addestrato, fino a svolgere un piccolo calcolo con i suoi pesi.

Capitolo 2 Come impara un modello

La parola «apprendimento» porta con sé immagini familiari: una persona che prova, si accorge di un errore, ricorda un'esperienza e cambia il proprio modo di agire. Quando la usiamo per una macchina, alcune di queste immagini aiutano e altre possono portarci fuori strada. Il modello cambia, ma dobbiamo capire che cosa cambi; migliora in alcuni compiti, ma dobbiamo sapere rispetto a quale obiettivo e su quali esempi il miglioramento venga osservato.



Figura 4 · Spazzare lo stesso sentiero: il gesto ripetuto richiama la pratica nel tempo, senza equipararla all'addestramento di una rete.

Immaginiamo un monaco che spazza ogni giorno lo stesso sentiero e accostiamo quel gesto ripetuto all'addestramento di una rete. La ripetizione suggerisce una disciplina, un percorso che prende forma nel tempo, e può prepararci a guardare l'addestramento senza considerarlo un atto magico. Il paragone si ferma però prima dell'intenzione: il monaco vive la propria pratica, mentre la rete viene regolata mediante dati, calcoli e un criterio scelto da chi la costruisce. Per comprendere il suo percorso, occorre rendere visibili questi passaggi.

Dai frutti a una previsione

Immaginiamo di voler distinguere mele e arance nelle fotografie. Una soluzione scritta interamente a mano potrebbe usare regole su colore e forma, ma troverebbe presto eccezioni: una mela può essere verde o rossa, un'arancia può apparire scura sotto una certa luce e lo sfondo può confondersi con il frutto. Il problema non è soltanto trovare una regola plausibile; è costruire un comportamento che continui a funzionare quando la foto cambia.

Una rete neurale offre un altro percorso. Le forniamo rappresentazioni numeriche delle immagini e un obiettivo, per esempio assegnare la categoria corretta. La rete trasforma quegli ingressi attraverso strati di unità collegate da pesi: ogni peso regola quanto un valore contribuisca a un calcolo successivo. Il risultato viene confrontato con l'esito desiderato, e quel confronto permette di modificare i parametri. Ripetendo il processo su molti esempi, possiamo ottenere una rete che distingue i frutti anche in fotografie non identiche a quelle incontrate.

La parola «possiamo» è necessaria, perché il successo non è automatico. Se quasi tutte le mele sono fotografate su un tavolo e le arance in una cassetta, il sistema può imparare una scorciatoia basata sullo sfondo. Avrà ridotto gli errori sugli esempi disponibili senza acquisire la capacità che ci interessava. Per accorgercene servono altre fotografie, tenute separate dall'addestramento e abbastanza varie da mettere alla prova ciò che il modello ha imparato.

Il medesimo ragionamento vale per la distinzione fra cani e gatti. Muso, occhi, orecchie e pelo offrono indizi che possono contribuire alla distinzione, ma nessuno è una definizione completa della categoria. La rete combina indizi distribuiti; noi dobbiamo controllare se quella combinazione rimanga utile con razze, pose e condizioni diverse. L'esempio ci insegna che imparare dagli esempi e riconoscere il problema non sono la stessa cosa: il criterio della prova deve collegarli.

Che cosa accade dentro una rete

Un neurone artificiale è una piccola unità di calcolo. Riceve alcuni valori numerici dell'immagine e assegna a ciascuno un peso, cioè un numero che ne regola il contributo. Moltiplica ogni ingresso per il suo peso e somma i risultati: questa è la somma pesata. Aggiunge poi un valore chiamato bias, che sposta la somma, e applica una funzione di attivazione, che trasforma il risultato prima di passarlo alle unità successive. Vediamo queste operazioni su due soli ingressi.

Una formula che possiamo leggere

La formula $y = \mathrm{ReLU}(Wx + b)$ descrive un possibile strato di rete, non l'intero funzionamento di un assistente. x è il vettore degli ingressi, cioè una lista ordinata di numeri; W è una matrice di pesi, in cui ogni riga combina quegli ingressi per una diversa uscita; b è un vettore di valori aggiunti, chiamati bias numerici. Il vettore y contiene i risultati. ReLU applica a ciascuno la regola $\mathrm{ReLU}(z) = \max(0, z)$: un valore negativo diventa zero, uno positivo resta uguale.

Immaginiamo un'uscita sola e due ingressi normalizzati, privi di unità fisiche, $x = (1; 0{,}5)$. I pesi sono $W = (2; -1)$ e il bias è $-0{,}5$. La somma è $2 \times 1 + (-1) \times 0{,}5 - 0{,}5 = 1$, e ReLU lascia 1. Se il secondo ingresso fosse 3, otterremmo $2 - 3 - 0{,}5 = -1{,}5$ e l'uscita diventerebbe 0. Il calcolo mostra come gli ingressi influenzino un risultato; non attribuisce al neurone una conoscenza della foto o della frase.

Se ci sono tre ingressi e due uscite, W deve avere due righe e tre colonne e b deve avere due valori: questa compatibilità permette di eseguire le somme. Nelle reti reali dimensioni, unità e normalizzazione dipendono dalla rappresentazione scelta, perciò non possono essere dedotte dal nome delle variabili. La formula aiuta a riconoscere la struttura; non sostituisce la descrizione del sistema in cui quella struttura viene usata.

Le unità sono organizzate in strati, e l'uscita di un passaggio diventa materiale per quello successivo. Il numero di strati e connessioni può essere grande, ma

la complessità complessiva non assegna a un singolo neurone una conoscenza del gatto o della mela. Le rappresentazioni utili sono distribuite nelle trasformazioni della rete. Come nel disegno di un giardino di sabbia, un elemento acquista il proprio ruolo attraverso le relazioni con gli altri; l'immagine aiuta a vedere l'insieme, senza sostituire le operazioni matematiche.

Per regolare i pesi dobbiamo confrontare la previsione con l'esito desiderato. Nelle fotografie dei frutti conosciamo la categoria indicata per ogni esempio e possiamo misurare quanto la risposta se ne discosti. Una funzione di perdita traduce questo confronto in un numero: durante l'addestramento cerchiamo di ridurlo. Resta da capire quali pesi cambiare e in quale direzione. Il calcolo seguente usa un obiettivo numerico ancora più semplice, così possiamo vedere un aggiornamento senza dover classificare un'intera fotografia.

Il criterio stesso ha dei limiti. Se misura la categoria del frutto, non misura necessariamente la robustezza rispetto alla luce; se misura la previsione di una sequenza, non misura direttamente la bontà etica di ogni uso della risposta. Il miglioramento di un numero è informativo soltanto quando sappiamo che cosa quel numero rappresenti e che cosa lasci fuori. Questo punto tornerà nei capitoli sulle capacità emergenti e sull'equità.

Un aggiornamento visto da vicino

Consideriamo un modello didattico molto piccolo: `y = w × x`. Il simbolo `x` indica l'ingresso, `w` il peso e `y` la previsione. Vogliamo che, per `x = 2`, il risultato sia `4`. Con il peso iniziale `w = 1`, otteniamo invece `y = 1 × 2 = 2`. Tutti questi valori sono numeri astratti, senza unità fisiche; questo esempio isola un aggiornamento numerico e non rappresenta da solo il classificatore dei frutti.

Misuriamo lo scarto con `L = (y - 4)^2`, dove `L` è la perdita e `4` il risultato desiderato. Sottraiamo il valore desiderato dalla previsione e poi eleviamo lo scarto al quadrato: `L = (2 - 4)^2 = 4`. Il quadrato rende positiva la perdita anche quando la previsione è troppo bassa. Se la previsione fosse esattamente `4`, lo scarto e la perdita sarebbero entrambi zero.

Per scegliere la direzione, proviamo prima due piccoli cambiamenti. Con w = 1,1, la previsione diventa 2,2 e la perdita 3,24; con w = 0,9, la previsione diventa 1,8 e la perdita 4,84. Vicino al punto di partenza, aumentare il peso aiuta e diminuirlo peggiora il risultato. Il gradiente descrive questa variazione locale senza dover provare separatamente ogni modifica.

Per la perdita scelta, il gradiente rispetto a w è 2 × (w × x - 4) × x. Il termine fra parentesi è lo scarto della previsione; il fattore 2 deriva dal quadrato e l'ultimo x tiene conto di quanto il peso contribuisce alla previsione. Con w = 1 e x = 2 otteniamo 2 × (2 - 4) × 2 = -8. Il segno negativo indica che, in quel punto, aumentando un poco w la perdita diminuisce.

Il passo di apprendimento stabilisce quanto ci muoviamo nella direzione suggerita. Con un passo di 0,1 sottraiamo al vecchio peso il gradiente moltiplicato per il passo: w_nuovo = 1 - 0,1 × (-8) = 1,8. La previsione diventa 1,8 × 2 = 3,6 e la perdita (3,6 - 4)^2 = 0,16. Il controllo conferma che questo aggiornamento ha ridotto la perdita da 4 a 0,16.

Un passo maggiore non garantisce un miglioramento maggiore. Ripartendo dal peso 1, con un passo di 0,5 avremmo w_nuovo = 1 - 0,5 × (-8) = 5: la previsione sarebbe 10 e la perdita 36. Avremmo superato troppo il valore utile. Il gradiente informa su una variazione locale; il passo deve essere scelto e il risultato controllato.

In una rete con molti pesi, la retropropagazione, o *backpropagation*, calcola i gradienti attraversando le operazioni in senso inverso rispetto alla previsione. L'ottimizzatore è l'algoritmo che usa quelle informazioni per decidere gli aggiornamenti. Dire che la rete «si corregge» riassume questi calcoli: il processo modifica valori secondo un criterio, senza riconoscere un errore come una persona. Non dobbiamo calcolare a mano milioni di gradienti per comprendere la distinzione fra misurare lo scarto, calcolare la direzione e aggiornare i pesi. Il successo del nostro esempio non garantisce che una rete reale impari sempre bene o trovi una soluzione unica.

Verificare lontano dagli esempi già visti

Per distinguere mele e arance, teniamo separate alcune foto dall'addestramento. Se la rete riconosce soltanto mele sullo stesso sfondo, può aver usato lo sfondo come scorciatoia. Proviamo frutti con luci, posizioni e sfondi diversi. La verifica deve incontrare il tipo di variazione che ci aspettiamo nell'uso.

La variante è un sistema che classifica messaggi. Se impariamo da messaggi tutti della stessa fonte, il successo potrebbe dipendere dal formato, non dal contenuto. Come cambieresti la prova? Includeresti messaggi di fonti diverse e controlleresti gli errori per categoria. Riusa il criterio quando «ha imparato» deve significare qualcosa di più di «funziona sugli esempi che gli abbiamo mostrato».

Due tempi diversi dell'apprendimento

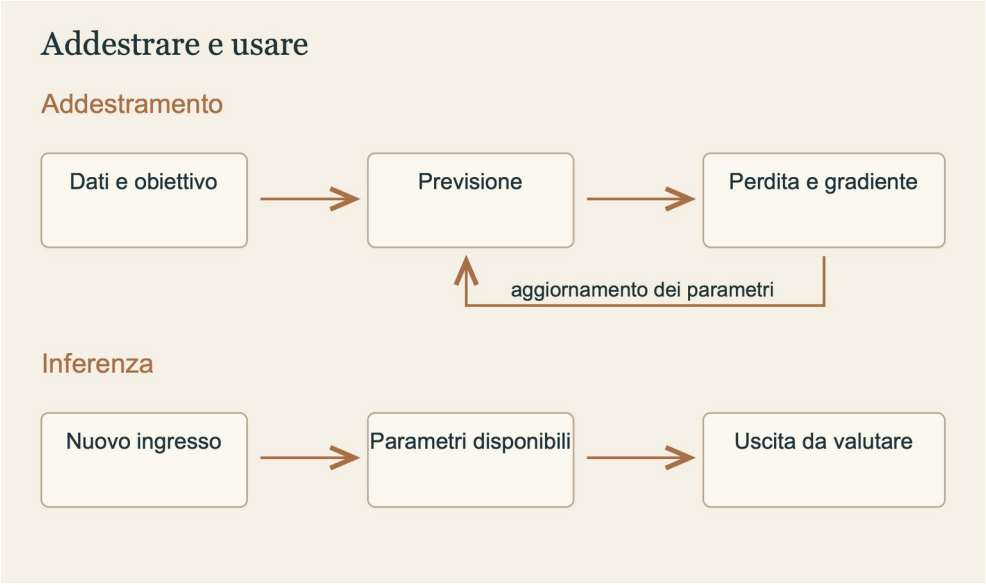


Figura 5 · Addestrare e usare una rete sono processi distinti

La figura separa l'addestramento dall'uso. Nel primo, dati e obiettivo permettono di calcolare una perdita e modificare parametri. Nel secondo, un nuovo ingresso viene elaborato con i parametri disponibili. Una correzione scritta

nella chat può entrare nel contesto della risposta successiva; non dimostra che i pesi siano stati riaddestrati durante la conversazione.

> **Generalizzare e memorizzare.** Generalizzare significa ottenere risultati utili su esempi non usati per regolare il modello. Memorizzare significa conservare nei parametri informazioni specifiche dei dati incontrati. Le due cose possono convivere. Un'ottima prestazione sugli esempi di addestramento non dimostra la generalizzazione, e la possibilità di memorizzazione ha conseguenze anche per i dati personali. [R2]

Dal riconoscere al generare

Fin qui la rete riceveva un'immagine e proponeva un'etichetta. Che cosa cambia se il compito diventa proseguire una sequenza, come «C'era una volta un gatto che…»? La struttura deve imparare relazioni che aiutino a produrre un elemento successivo, anziché soltanto a scegliere una categoria. È uno dei modi in cui una rete può diventare un modello generativo: l'uscita entra nella costruzione di un nuovo contenuto.

Nel preaddestramento di molti modelli di linguaggio, l'obiettivo è prevedere il token successivo. Il compito è compatto nella formulazione, ma il linguaggio che lo rende possibile è ricco: contiene fatti, narrazioni, argomentazioni, dialoghi, procedure e stili. Per prevederlo bene, il modello sviluppa rappresentazioni che possono poi essere usate in modi diversi. Una richiesta può orientare queste rappresentazioni verso una poesia, un messaggio, una spiegazione o del codice, senza che ciascuna risposta sia stata scritta in anticipo.

L'immagine del musicista che improvvisa entro un raga mette in evidenza una relazione fra libertà e struttura. Il modello non improvvisa attraverso un'esperienza musicale, ma produce alternative entro possibilità apprese. Le combinazioni possono essere nuove e utili; non dobbiamo per questo attribuire loro un'origine indipendente dai dati o garantire che ogni combinazione abbia un fondamento nel mondo.

Il rapporto fra forma e significato può suggerire un accostamento al richiamo buddhista a forma e vuoto. L'analogia apre una domanda, ma non offre una

spiegazione tecnica: il vuoto di una tradizione spirituale non coincide con un meccanismo statistico. Le rappresentazioni del modello si sviluppano attraverso l'addestramento; il rapporto umano con il significato di una frase appartiene anche alla situazione in cui la frase viene letta e usata.

L'illusione che nasce dalla continuità

Quando un chatbot risponde con un tono preciso e una struttura convincente, può sembrare che abbia seguito il nostro pensiero come farebbe una persona. In alcuni compiti produce davvero una soluzione valida; in altri riproduce la forma di una soluzione senza rispettarne i passaggi. Uno studente che legge una spiegazione di fisica, un decisore che usa un riassunto giuridico e una persona che cerca ascolto incontrano tutti il rischio di estendere al sistema la fiducia suscitata dalla voce.

Nel primo caso occorre controllare il procedimento e le ipotesi; nel secondo confrontare il riassunto con il documento e con una valutazione competente; nel terzo riconoscere che la sensazione di essere ascoltati non dimostra una partecipazione emotiva reciproca. Il modello può svolgere un lavoro significativo in tutte queste situazioni, ma non lo stesso lavoro che immaginiamo quando vediamo soltanto la continuità della conversazione.

> **Il limite della non-mente.** Il richiamo a *muga-mushin*, spesso reso con «non-mente», invita a interrogare un'azione che sembra priva di un sé. Espressioni come «non-mente» appartengono però a pratiche e tradizioni umane e non indicano semplicemente l'assenza di pensiero o di ego. L'analogia può farci riflettere su ciò che proiettiamo nella macchina; non le assegna un percorso o una realizzazione spirituale.

Anche il ricordo va distinto dalla continuità visibile nella chat. Un servizio può conservare messaggi e renderli nuovamente disponibili, oppure recuperare preferenze da una memoria esterna. Il modello può usarli come parte dell'ingresso della nuova esecuzione. Questa organizzazione aiuta a mantenere il filo, ma è diversa dal ricordo autobiografico con cui una persona collega esperienze e identità.

Il contesto come maestro silenzioso

Per scegliere una continuazione, un modello generativo usa le informazioni disponibili nel contesto. La frase «Il filosofo sedeva in silenzio perché…» può proseguire con «aveva detto abbastanza», «la verità era oltre le parole» o «non aveva campo»: la stessa apertura può condurre a una riflessione, a un'immagine poetica o a una battuta, secondo il testo e le istruzioni che la accompagnano. Non basta dunque chiedere quale risposta sia più probabile in astratto; occorre guardare quale spazio di possibilità abbiamo costruito.

La finestra di contesto delimita la quantità di token gestibile nelle condizioni del modello e del servizio. Il limite si esprime in token e varia da un sistema all'altro; per un uso concreto occorre controllare quello dichiarato dal servizio e la parte disponibile per documenti, istruzioni e risposta. Quando il limite non permette di includere tutto, il servizio può escludere parti, riassumerle o recuperarle selettivamente. La perdita non avviene necessariamente togliendo sempre il testo più vecchio, e una finestra grande non garantisce l'uso corretto di ogni dettaglio.

Questa distinzione ha una conseguenza pratica. Se una risposta trascura una condizione importante, possiamo controllare se la condizione fosse davvero disponibile nell'esecuzione, se fosse leggibile e se la richiesta la rendesse rilevante. Aggiungere più testo non è sempre la soluzione: talvolta serve selezionare meglio, rendere esplicito un vincolo o dividere il lavoro in passaggi che possiamo verificare. Il contesto non rende il sistema infallibile, ma permette di capire una parte delle ragioni del risultato.

Percorrere senza possedere

L'immagine del monaco che copia un sutra e del rastrello che traccia la sabbia porta l'attenzione sulla forma che viene trasmessa. Un modello riprende tracce dell'espressione umana e le ricombina; chi lo usa può osservare in quelle tracce una procedura, una conoscenza, un pregiudizio o una possibilità nuova.

Il lavoro dell'attenzione consiste nel distinguere queste presenze, anziché assegnare loro in anticipo il nome di saggezza.

Ora la parola «imparare» ha un percorso riconoscibile: produrre una previsione, misurare lo scarto, aggiornare i parametri e provare il risultato su esempi nuovi. Abbiamo anche distinto quel percorso dall'uso del contesto nella conversazione. Il passo successivo è l'interazione: come formulare una richiesta, quali informazioni offrirle e come riconoscere ciò che ha cambiato nella risposta.

Capitolo 3 Quando il linguaggio diventa interfaccia

Una domanda può cercare un'informazione e può anche rendere più chiaro ciò che non avevamo ancora formulato. Succede quando proviamo a spiegare un problema a un'altra persona e, nel costruire la frase, scopriamo una condizione lasciata implicita o una scelta già compiuta senza accorgercene. L'incontro con un modello di linguaggio offre una possibilità simile, anche se la relazione è diversa: scriviamo un'intenzione e riceviamo un testo che ci aiuta a vedere quale forma abbia preso.

Nel capitolo precedente abbiamo seguito l'addestramento e il ruolo del contesto. Ora entriamo nella richiesta, il punto in cui quei meccanismi diventano esperienza d'uso. Il modello non riceve soltanto un argomento: riceve parole che definiscono un pubblico, un tono, un compito e una direzione. «Spiega questo testo», «Aiutami a scrivere a qualcuno che ho deluso» e «Rendi più semplice questo passaggio» possono chiamare capacità vicine, ma chiedono lavori differenti e devono essere giudicati con criteri differenti.

È questo il cambiamento dell'interfaccia linguistica. Il linguaggio naturale era già stato usato nell'informatica prima degli strumenti attuali, ma la versatilità e la diffusione dei sistemi contemporanei lo portano al centro di molte attività. Dove dovevamo cercare un comando o imparare una sintassi, possiamo ora descrivere un bisogno. Questa facilità apre possibilità e, nello stesso movimento, può nascondere vincoli: una frase quotidiana sembra trasparente anche quando lascia molto da decidere al sistema.

Una richiesta non è una procedura completa

Consideriamo una regola scritta in un programma:

```
if temperature > 100:
    activate_cooling()
```

L'esempio è incompleto come sistema reale, ma permette di vedere una distinzione: la condizione e l'azione sono espresse secondo una sintassi, e la loro esecuzione dipende dal programma e dai valori ricevuti. Una richiesta in linguaggio naturale, invece, lascia spazio a interpretazioni apprese. «Suggerisci cinque ricette vegetariane con zucchine e quinoa, semplici e adatte a un principiante» indica ingredienti, quantità, destinatario e stile; non definisce ogni operazione che il modello seguirà né garantisce la qualità culinaria di ogni proposta.

Se aggiungiamo «Sei uno chef professionista», orientiamo il registro e il tipo di contenuto che il sistema può produrre. Non gli conferiamo una qualifica reale. La semplicità può influenzare il vocabolario, la lunghezza dei passaggi e la scelta degli ingredienti; la richiesta di vegetarianismo delimita un vincolo che dovremo comunque rileggere nelle ricette. Il ruolo aiuta a dare una forma alla risposta, mentre la correttezza rimane da verificare.

Possiamo accostare questa interazione alla risonanza di un diapason vicino a una corda. La nota non nasce da una conversazione umana, ma da una struttura che risponde a un ingresso; in modo analogo, una richiesta può far emergere schemi linguistici coerenti con il tono che abbiamo offerto. L'immagine chiarisce l'orientamento, purché non ci faccia credere che il modello riceva un'esperienza o un'intenzione come le riceve una persona.

Si parla talvolta di «programmazione in prosa» per descrivere il prompt. È una formula suggestiva, ma va usata con prudenza: la chiarezza della richiesta contribuisce al risultato insieme a dati, addestramento successivo, contesto, strumenti e regole del servizio. Non tutto ciò che accade è contenuto nelle nostre parole, e una richiesta precisa non rende inevitabile una risposta corretta.

Un prompt preciso deve incontrare una prova

Prendiamo un'altra richiesta, questa volta su una tabella trimestrale. Supponiamo che i ricavi siano 100, 120, 90 e 130, tutti in migliaia di euro. Chiediamo di descrivere l'andamento senza inventare cause. Una risposta adeguata

segnala il calo da 120 a 90 nel terzo trimestre e la risalita a 130 nel quarto. Il calo è 120 - 90 = 30 migliaia di euro. Per esprimerlo rispetto al trimestre precedente, dividiamo 30 per 120 e moltiplichiamo per 100: 30 / 120 × 100 = 25%. Il valore 120 è la base del confronto, non il totale dell'anno. I quattro valori, da soli, non spiegano il motivo del calo.

Se il modello attribuisce il calo a una campagna fallita, ha oltrepassato i dati. La verifica consiste nel confrontare ogni frase con la tabella e rifare la percentuale. Una richiesta di tono più elegante non corregge quel salto. Chiediamo invece di separare l'osservazione dalle ipotesi e di indicare quali dati servirebbero per provarle.

Nella variante i valori rappresentano presenze, non ricavi. La percentuale si calcola allo stesso modo, ma cambia la spiegazione utile e cambiano i dati necessari per cercare una causa. Riusa questo procedimento quando il linguaggio deve raccontare numeri: unità, base del confronto e distinzione tra andamento e causa devono restare visibili anche se cambi registro.

Una sosta nel giardino del Transformer

Per capire come il modello colleghi parti del testo, seguiamo una frase: «I sat by the river bank», cioè «Sedevo sulla riva del fiume». La parola inglese *bank* può indicare anche una banca; *river* aiuta a capire quale uso sia pertinente. Vedremo come la frase diventa una sequenza di numeri, come le sue parti vengono messe in relazione e come il generatore prepara la continuazione. Possiamo immaginare questa sosta come l'osservazione di un giardino: da lontano vediamo un disegno continuo, mentre avvicinandoci distinguiamo i passaggi che lo rendono possibile.

Le parole che offriamo

La tokenizzazione trasforma il testo in una sequenza di identificatori. I token sono unità di elaborazione, non necessariamente parole o sillabe e non necessariamente unità con un significato autonomo. Per «I am learning» possiamo immaginare una divisione come ["I", " am", " learn", "ing"]: è un esempio illu-

strativo, perché la divisione effettiva dipende dal vocabolario del sistema. Gli spazi possono far parte delle unità, e una parola italiana può essere divisa diversamente da una inglese.

Nella frase «I sat by the river bank» possiamo immaginare, per esempio, le unità I, sat, by, the, river, bank. Ogni unità riceve un identificatore nel vocabolario del sistema. È ancora una divisione illustrativa: un altro vocabolario potrebbe spezzare diversamente le parole. Il passo successivo parte da quegli identificatori e li trasforma in rappresentazioni numeriche.



Figura 6 · Dal flusso ai frammenti: i ciottoli raccolti evocano unità distinte; il significato del testo dipende dalle relazioni fra i token.

L'immagine dei ciottoli raccolti da un torrente rende visibile il passaggio dal flusso ai frammenti. Ci aiuta anche a riconoscere il limite dell'analogia: un token non è già un piccolo pensiero. Il significato utile del testo emerge dalle rappresentazioni e dalle relazioni che il modello costruisce, non dalla pretesa di assegnare a ogni frammento una comprensione isolata.

La forma numerica del linguaggio

Dopo la tokenizzazione, ogni identificatore viene associato a una lista ordinata di numeri, chiamata vettore. Questa rappresentazione numerica è un embed-

ding. L'ordine conta: la prima coordinata si confronta con la prima di un altro vettore, la seconda con la seconda, e così via. Possiamo immaginare due coordinate come un punto su un foglio; il modello ne usa molte di più, in uno spazio matematico che non possiamo disegnare interamente.

Torniamo a *bank*. La rappresentazione iniziale del suo token non basta da sola a decidere se stiamo parlando di un fiume o di denaro. Le trasformazioni successive la metteranno in relazione con gli altri token della frase. Negli embedding di parole, relazioni apprese possono rendere vicini termini usati in modi simili. Il lavoro di Mikolov e colleghi mostra come i vettori possano sostenere confronti sintattici e semantici. [R11] Quelle relazioni descrivono regolarità dei materiali e del modello, non una mappa definitiva del significato.

Coppie come «re» e «regina» o «medico» e «infermiere», insieme a parole come «giustizia», «equità», «legge» e «potere», ci aiutano a immaginare una geografia del linguaggio. Non sono distanze universali: dipendono dai dati, dal modello e dal criterio di confronto. Anche una vicinanza statisticamente utile può contenere uno stereotipo, perciò lo spazio numerico non certifica come le persone debbano essere descritte o trattate.

Leggere un'analogia vettoriale. In re - uomo + donna ≈ regina, i nomi indicano vettori. Per vedere come si eseguono le operazioni, inventiamo tre liste di due numeri: (4; 2) - (1; 0) + (0; 1) = (3; 3). La prima coordinata diventa 4 - 1 + 0 = 3, la seconda 2 - 0 + 1 = 3. Questi valori senza unità sono soltanto didattici e non misure reali delle parole. Nell'analogia, ≈ significa che il vettore ottenuto è vicino a quello di «regina» secondo una misura scelta; non significa uguaglianza. Occorre controllare quella vicinanza nei vettori effettivi. L'analogia non vale per ogni modello o per tutti gli embedding contestuali e non definisce socialmente persone o ruoli.

Le relazioni di attenzione

Nella nostra frase, *river* offre un indizio utile per la rappresentazione di *bank*. La self-attention, o attenzione interna alla sequenza, permette di combinare informazioni delle posizioni disponibili dando loro contributi diversi. I pesi vengono calcolati a partire dal contenuto delle rappresentazioni. «Attenzione»

è il nome di questo meccanismo matematico; non indica che il modello sia cosciente del fiume.

Per leggere una combinazione, immaginiamo due vettori già trasformati: (2; 0) e (0; 4). Usiamo pesi illustrativi 0,75 e 0,25, che sommano a 1. Moltiplicando ogni coordinata per il proprio peso otteniamo (1,5; 0) e (0; 1); sommando le coordinate corrispondenti, il risultato è (1,5; 1). Il primo vettore contribuisce di più, ma il secondo rimane presente. Questi numeri astratti mostrano una somma pesata; non sono i valori reali di *river* e *bank* e non ricostruiscono un intero strato di attenzione.

Un generatore autoregressivo aggiunge un token alla volta usando la sequenza disponibile. Durante il calcolo, una maschera causale impedisce a una posizione di usare token successivi: il testo già presente può contribuire, una continuazione futura non può essere anticipata. [R1] La parola «maschera» indica quindi un vincolo su quali collegamenti sono ammessi. La combinazione delle informazioni e il vincolo sul futuro svolgono due lavori distinti.

Se cambiamo *river bank* in *investment bank*, cambiano gli indizi utili: *investment* orienta la rappresentazione verso un contesto finanziario. Questo confronto chiarisce perché non possiamo assegnare a un token un significato completo e immutabile. Possiamo poi osservare la continuazione, senza scambiare la capacità di collegare parole per una comprensione umana della scena.

Transformer e modelli diversi. Vaswani e colleghi introdussero il Transformer nel lavoro *Attention Is All You Need* del 2017. [R1] Le reti ricorrenti, dette RNN, trasmettono uno stato lungo la sequenza; le LSTM, sigla di *long short-term memory*, sono una loro variante progettata per mantenere informazioni attraverso più passaggi. Il Transformer usa meccanismi di attenzione e altri strati di trasformazione. Non tutti i suoi usi seguono il percorso di un generatore: BERT, presentato da Devlin e colleghi, costruisce rappresentazioni utilizzando contesto sia a sinistra sia a destra. [R12] Il confronto spiega una differenza di funzionamento, senza stabilire che un'architettura sia adatta a ogni compito.

Il testo che continua

Gli strati successivi trasformano le rappresentazioni e combinano informazioni distribuite. Non esiste una sequenza universale in cui il primo strato si occupi soltanto di grammatica e l'ultimo soltanto di intenzione. I calcoli interagiscono; una singola freccia di attenzione non costituisce una spiegazione completa del motivo per cui una risposta è stata prodotta.

Dopo le trasformazioni, il generatore calcola probabilità per il token successivo. A «I sat by the river bank» potrebbe aggiungere una continuazione come «and», poi usare la sequenza aggiornata per un nuovo passo. È una possibilità illustrativa, non una risposta misurata. Ogni token selezionato entra nel contesto del calcolo seguente. Se riceve «The quick brown fox jumps over the…», può proporre «lazy» senza dover visualizzare un cane; la scelta concreta dipende anche dalla procedura di selezione. Il parallelismo aiuta a elaborare più posizioni nell'addestramento e nella lettura iniziale, mentre la produzione autoregressiva rimane legata a ciò che è stato già generato.

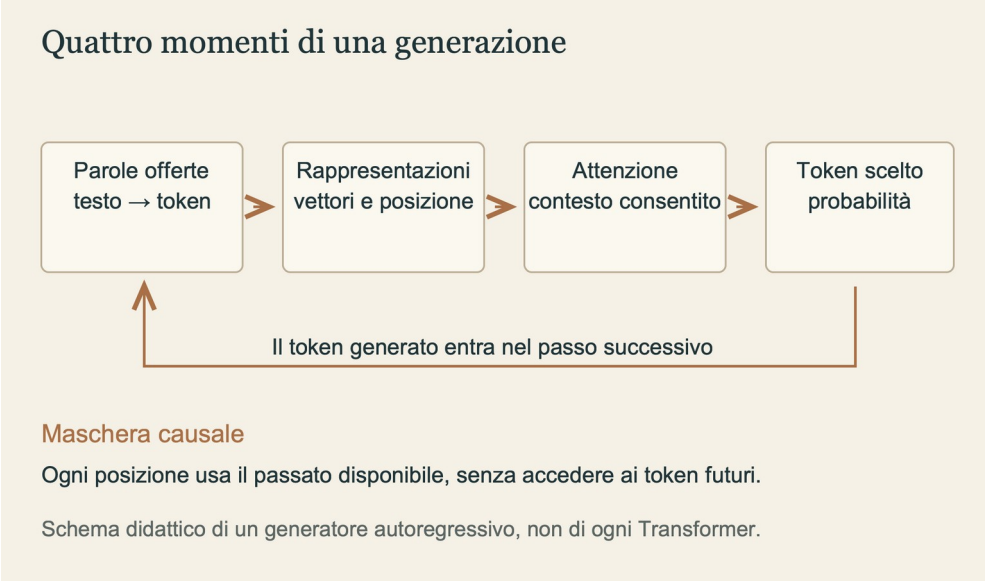


Figura 7 · Quattro momenti di una generazione

La figura rende visibili quattro passaggi del processo: le parole offerte, la forma del significato, l'attenzione interna e la parola che emerge. Il ritorno

alla sequenza rende esplicito il ciclo. Possiamo vedere come venga costruita una continuazione senza attribuire al processo il silenzio, l'intenzione o la meditazione di una persona.

La domanda rivela chi sta domandando

Proviamo quattro richieste sull'evoluzione: «Spiega l'evoluzione», «Spiegala come se avessi dieci anni», «Scrivi un saggio persuasivo a favore della teoria» e «Immagina Charles Darwin che la difende davanti a un vittoriano scettico». L'argomento rimane vicino, ma il lavoro cambia: spiegazione generale, adattamento didattico, argomentazione e scena letteraria. L'ultima risposta non va presentata come una citazione storica di Darwin, perché il ruolo richiesto è una simulazione.

Formulare queste differenze ci costringe a osservare le nostre aspettative. Vogliamo capire, persuadere, trovare una metafora o ricevere conferma? Quale conoscenza attribuiamo al lettore? Quali supposizioni abbiamo incorporato nella domanda? Questo lavoro è metacognitivo: riguarda il modo in cui pensiamo al nostro pensiero. Il prompt diventa utile anche quando ci restituisce un limite che non avevamo ancora visto.

Il richiamo al kōan ci invita a una riflessione personale: una domanda può modificare lo sguardo di chi la porta, oltre a chiedere una risposta. Nelle tradizioni Zen il kōan appartiene a pratiche e percorsi specifici; non lo riduciamo a una tecnica per scrivere prompt. Ne conserviamo l'invito a non considerare conclusa la domanda soltanto perché la schermata ha prodotto parole.

Consideriamo tre contesti molto diversi. Per un assistente di compagnia aerea, «Rispondi con calma e offri soluzioni» orienta il tono ma non autorizza a inventare un rimborso: servono informazioni e regole reali della compagnia. Per spiegare un tema scientifico a bambini attraverso un gufo narratore, la personificazione può aiutare il coinvolgimento ma non sostituisce la revisione della correttezza. Per descrivere una tabella trimestrale, invece, il sistema deve conservare valori, unità e confronti prima di scegliere una voce elegante.

Queste distinzioni mostrano che un prompt ben formulato non è una formula buona per ogni situazione. Delimita il compito e rende possibile una verifica adatta. Se chiediamo più semplicità, controlliamo che il meccanismo essenziale resti; se chiediamo più cortesia, controlliamo che l'impegno non cambi; se chiediamo una lettura dei dati, controlliamo che un'osservazione non diventi una causa inventata.

Una voce porta con sé una cornice

Chiedere a un modello di parlare come un politico, un mago o un collega passivo-aggressivo rende visibile il legame fra stile e aspettative. Il risultato può riprendere modi di argomentare, parole e ritmo associati al ruolo. Non c'è una qualifica reale dietro la maschera, e il ruolo non garantisce la qualità del contenuto: una voce accademica può citare una fonte inesistente e una voce gentile può nascondere una condizione importante.

Possiamo pensare alle parole come a tracce che ci orientano nella ricerca della verità: il linguaggio può indicarci un problema e orientarci verso un riscontro, ma il modo in cui lo presenta non è già quel riscontro. Una frase sicura può essere falsa; una prudente può essere corretta o evasiva. Leggere la risposta significa quindi osservare ciò che dice e la cornice che rende naturale dirlo in quel modo.

I dati di addestramento possono includere libri, manuali, siti, conversazioni, poesie, documenti legali e molti altri materiali, con qualità e prospettive diverse. Non tutti i modelli hanno incontrato le stesse fonti, né l'intero linguaggio umano è disponibile in forma digitale. Le voci dominanti nei materiali possono influenzare ciò che appare più familiare nella risposta. Il capitolo sullo specchio culturale svilupperà questo limite; qui possiamo già riconoscerlo nel pubblico che ogni tono sembra immaginare.

Se chiediamo «Spiega come se avessi cinque anni», abbiamo introdotto un'idea di competenza e di cura, oltre alla semplicità. Se chiediamo «Convincimi che ho ragione», abbiamo orientato il compito verso la conferma. Possiamo cambiare richiesta e osservare quale parte del risultato cambia: è un

modo di leggere anche noi stessi, purché non presumiamo che ogni variazione del testo riveli qualcosa di vero sulla nostra identità.

Registri linguistici e adattamento al contesto

Il modello può cambiare tono e registro perché ha appreso molte forme d'uso del linguaggio. La varietà dipende dai materiali e dall'addestramento, e non comprende necessariamente ogni contesto umano. La tabella offre cinque estratti illustrativi sull'arrivo della stampa a caratteri mobili nell'Europa di Gutenberg. Confrontiamo il modo di presentare il tema, senza prendere ciascun estratto come una ricostruzione completa della storia della stampa.

Registro	Estratto della risposta
Elevato o cerimoniale	«Onoriamo il momento sacro in cui Gutenberg donò all'umanità la sua voce.»
Accademico o tecnico	«La stampa, sviluppata da Gutenberg nel XV secolo, rivoluzionò…»
Neutro o generale	«La stampa rese più rapida ed economica la produzione dei libri, cambiando la storia.»
Colloquiale o familiare	«Prima della stampa, copiare libri era una gran fatica. Poi arrivò Gutenberg.»
Gergo informale	«Gutenberg tirò fuori il copia-incolla originale: da lì i libri spopolarono.»

Ogni registro rivela valori e supposizioni diversi. Spiegare l'evoluzione, spiegarla a un bambino e difenderla con la voce di Darwin sono cornici diverse della stessa conoscenza. Il linguaggio è ciò che diciamo e il modo in cui lo collochiamo. Con un prompt diamo forma al significato attraverso registro, tono e contesto.

> «Il tono riprende forme apprese e può ricombinarle: la familiarità di una voce non ne certifica il contenuto.»

Questa consapevolezza aiuta a interagire criticamente, progettare domande migliori e leggere con più cura le risposte.

La stessa domanda con una voce diversa

> «Non chiedere al modello di dirti che cosa sia vero. Chiedigli di mostrarti come stai domandando.» Riflessione proposta nel libro.

Anche conservando il contenuto, lo stile può trasformare radicalmente la risposta. **Tono scientifico:** «Spiega il cambiamento climatico in modo scientificamente accurato». Risposta: «Il cambiamento climatico riguarda variazioni a lungo termine della temperatura globale, delle precipitazioni e della composizione atmosferica, dovute in gran parte a emissioni antropiche di gas serra come CO_2 e metano. Questi cambiamenti alterano ecosistemi, correnti oceaniche e sistemi meteorologici».

Tono poetico: «Spiega il cambiamento climatico come un poeta». Risposta: «La Terra ha la febbre e sussurra tempeste e silenzio. I ghiacciai piangono nei mari che salgono e il cielo porta il peso di un respiro di carbonio. Ciò che era ritmo è ora frattura. Le stagioni non dormono più come un tempo».

Saggezza di una persona anziana: «Spiegalo come un nonno che parla a un bambino». Risposta: «Sai come le stagioni vanno e vengono, tesoro? Ultimamente il tempo è confuso. È come se Madre Natura non riuscisse a riposare abbastanza. Abbiamo usato troppo, bruciato troppo; ora è un po' stanca e cerca di dirci qualcosa».

Narratore eroico: «Spiegalo come un narratore di fumetti». Risposta: «In un mondo un tempo in equilibrio fra ghiaccio e fuoco, sorse un nemico silenzioso: il carbonio. Invisibile e instancabile si insinuò nei cieli, rafforzando l'effetto serra. Ora i più grandi difensori del pianeta, scienza, azione e determinazione umana, devono reagire prima che il tempo finisca!».

Questi estratti illustrativi mostrano come cambi lo stile e come possa cambiare anche il contenuto: febbre, stanchezza e lotta sono metafore, non spiegazioni fisiche. Per conservare il meccanismo occorre tornare alla descrizione dei gas

serra e verificare che la riscrittura non lo alteri. [R13] Il tono modifica coinvolgimento, empatia e autorità. Chiedendo informazioni definiamo anche come presentarle.

Il punto in cui la voce non basta

La fluidità riduce l'attrito e può rendere un compito accessibile, ma rende meno visibili i passaggi che non sono stati verificati. Un modello può dichiarare di non sapere, rifiutare una richiesta o esprimere incertezza; questi comportamenti dipendono anche dall'addestramento e dalle regole del servizio. Nessuno ci esonera dal controllo di una risposta che verrà usata come fatto, consiglio o decisione.

Possiamo chiamare «limite del vuoto» il punto in cui il linguaggio prosegue senza un terreno adeguato. Può essere una spiegazione basata su informazioni superate, una figura storica inventata o una conclusione etica che ignora le persone coinvolte. L'immagine non descrive un vuoto spirituale nel modello: ci invita a vedere il limite del rapporto fra parole e riscontro. Quando lo riconosciamo, possiamo cercare una fonte, delimitare il compito o lasciare aperta la risposta.

Nella scena letteraria dello studente che chiede la mente dell'illuminazione e riceve «Non-mente», il silenzio successivo appartiene a una pratica umana. Un generatore può produrre parole anche intorno a «Qual era il tuo volto originario prima di nascere?», «Un cane ha la natura di Buddha?» o «Qual è il suono di una sola mano?». La presenza di quelle parole non dimostra che abbia abitato la domanda. L'accostamento vale se ci aiuta a distinguere la generazione dal modo in cui scegliamo di vivere e portare una risposta.



Figura 8 · La tazza e la pausa: lasciare spazio significa leggere la risposta, confrontarla con le fonti e tornare alla domanda prima di proseguire.

L'immagine della tazza ci è utile perché lascia spazio. Anche una conversazione con il modello può lasciare uno spazio se non la usiamo soltanto per riempire ogni pausa: possiamo leggere, controllare e tornare alla domanda. La rivoluzione dell'interfaccia linguistica acquista così un significato più ampio della velocità, perché ci chiede di imparare a riconoscere quale lavoro abbiamo affidato alle parole.

Il modello può offrirci uno specchio parziale e una possibilità espressiva. La scelta di farne un'autorità, una bozza o un'occasione di riflessione rimane dentro il nostro uso e dentro il progetto del servizio. Il confronto con la tabella trimestrale ci ha dato un criterio concreto: la voce può cambiare, ma i numeri e i loro limiti devono restare riconoscibili. Portiamo questo criterio nel capitolo successivo, dove una soluzione riuscita ci chiederà di distinguere la capacità osservata dal modo in cui la misuriamo.

Parte II Tra capacità e illusioni



La piuma e le tracce d'inchiostro aprono una sosta fra ciò che chiediamo e ciò che riceviamo. La piuma suggerisce leggerezza e disponibilità al cambiamento; le tracce richiamano le forme che una risposta lascia sulla pagina. Sono immagini per orientare la nostra attenzione: non attribuiscono alla macchina una realizzazione spirituale.

Dopo aver visto come una richiesta entri nel sistema, possiamo osservare ciò che ne emerge. Una risposta può risolvere un problema, creare una storia o collegare idee lontane; può anche inventare un riferimento o riprodurre una distorsione culturale. La stessa fluidità attraversa questi esiti e non ci permette, da sola, di distinguerli.

Questa parte tiene insieme capacità e limiti. Una soluzione sui treni ci permetterà di controllare un calcolo; una citazione plausibile ci chiederà di trovare e leggere la fonte; una traduzione ci mostrerà dove il testo aggiunge informazioni sulle persone. Da questi casi risaliremo ai criteri con cui giudichiamo riuscita, attendibilità ed equità. La leggerezza della piuma ci invita a cambiare prospettiva quando il controllo mostra che la prima impressione non basta.

Capitolo 4 Le capacità che emergono

Dopo aver seguito il linguaggio come interfaccia, possiamo tornare alla sorpresa che spesso apre una conversazione con un modello: una risposta sembra andare oltre il compito semplice con cui immaginavamo fosse stato addestrato. Offre un'analogia, risolve un problema o conserva uno stile mentre cambia argomento. Per comprendere questa sorpresa dobbiamo distinguere la capacità osservata, le condizioni che la rendono possibile e il modo in cui la misuriamo.

Possiamo accostare la comparsa di una capacità al momento in cui una lunga pratica rende naturale un gesto prima difficile. L'immagine richiama il ruolo della preparazione e delle condizioni, ma non identifica l'addestramento con l'esperienza di chi pratica lo Zen. Una rete può migliorare senza vivere quel percorso; la domanda tecnica riguarda il comportamento che possiamo osservare e il criterio con cui lo riconosciamo.

Immaginiamo una prova didattica che richieda quattro passaggi corretti. Un sistema ne svolge due, poi tre, poi quattro. Se assegniamo un punto soltanto alla risposta interamente corretta, i primi due tentativi valgono zero e il terzo vale uno: il grafico sembra mostrare un salto. Se contiamo invece i passaggi corretti, vediamo 2, 3 e 4, un progresso graduale. Sono numeri inventati per leggere la differenza fra le due misure, non risultati di un esperimento.

> **Perché guardare il punteggio.** Schaeffer, Miranda e Koyejo hanno studiato come alcune metriche possano produrre salti apparenti nelle prestazioni. [R4] Il lavoro invita a esaminare la misura prima di attribuire il salto a una nuova capacità. Non permette di concludere che ogni forma di emergenza sia illusoria: servono compiti, confronti e dati dichiarati.

Il preaddestramento di molti modelli generativi usa la previsione del token successivo. È un obiettivo compatto, ma impararlo può richiedere rappresentazioni complesse delle relazioni presenti nei testi. L'addestramento successivo

può sviluppare esplicitamente l'esecuzione di istruzioni: non tutte le capacità derivano da un unico passaggio. [R3] Il comportamento osservato dipende quindi dai dati, dall'architettura, dalle risorse di addestramento e dalle prove con cui lo valutiamo. Una continuazione di testo può prendere la forma di un'analogia o di una soluzione; dobbiamo ancora stabilire quanto quella soluzione sia affidabile e trasferibile.

Prevedere e seguire istruzioni. Ouyang e colleghi studiano come ulteriori fasi di addestramento possano orientare un modello a seguire le richieste delle persone. [R3] Includono esempi di risposte e valutazioni umane: cambia l'obiettivo rispetto alla sola previsione del token successivo. Il lavoro chiarisce perché una capacità d'uso possa dipendere da più fasi. Non garantisce che tutte le istruzioni vengano seguite correttamente o che il giudizio umano usato nell'addestramento rappresenti ogni persona.

Qui nasce la domanda sul comportamento emergente: una capacità nuova è comparsa, oppure abbiamo cominciato a riconoscere un miglioramento che prima non superava il nostro criterio di successo? Quando un mucchio di sabbia diventa una duna, e quando una goccia d'inchiostro diventa un dipinto? Le immagini aiutano a sentire la soglia, ma non la misurano. Nei modelli, una prova superata a una certa scala e fallita a una scala minore può suggerire una discontinuità; per comprenderla occorre guardare anche agli errori intermedi e alle condizioni del confronto.

Gli offri un indovinello e può risolverlo; chiedi codice e può scriverlo con commenti. Anche una domanda medica può ricevere una risposta dal linguaggio professionale, la cui coerenza non certifica l'accuratezza clinica. L'effetto è sorprendente perché riconosciamo forme del nostro ragionamento. La risposta, però, non dimostra consapevolezza, curiosità o intenzione: nasce da rappresentazioni apprese e dai passaggi di addestramento del sistema. La distinzione ci permette di apprezzare una capacità senza affidarle un'autorità che la sola forma del testo non può sostenere.

Nel linguaggio dell'eco, potremmo immaginare uno sconosciuto che ha incontrato un numero immenso di conversazioni e ricombina i modi in cui le persone hanno risposto. L'immagine non descrive un archivio di conversa-

zioni memorizzate e recuperate una per una: i parametri rappresentano regolarità apprese, pur potendo conservare anche frammenti specifici. L'eco si adatta al contesto presente e può risultare utile a chi la riceve, senza che la somiglianza con una risposta umana dimostri un'esperienza condivisa.

La scala può contribuire al risultato, insieme alla qualità e varietà dei dati e alle scelte di addestramento; aumentare gli strati o i parametri non garantisce da solo una particolare capacità. La fluidità che osserviamo può ricordare il momento in cui una lunga pratica rende naturale un gesto. Nell'immagine del sutra, chi recita comincia a sentirne il ritmo dall'interno; nel modello, invece, stiamo osservando un comportamento che dobbiamo descrivere e mettere alla prova. L'analogia conserva la sorpresa e ci lascia il compito della verifica.

Si parla di capacità emergente quando un comportamento riconoscibile a una scala maggiore non era evidente a una scala minore. La difficoltà è stabilire se sia cambiato bruscamente il comportamento oppure il nostro giudizio di successo. L'immagine di una soglia fra dieci e tredici miliardi di parametri resta illustrativa, senza indicare una legge o un risultato misurato. Una metrica, cioè la regola con cui assegniamo il punteggio, può rendere improvviso un miglioramento graduale. [R4]

Il modello può produrre passaggi di ragionamento, rispondere a «Perché?» e «Che cosa accadrebbe se…?» e interpretare ruoli appresi dai testi. Descriverli come echi della cognizione umana ci ricorda la loro relazione con i nostri materiali. Per sapere che cosa sappia fare il sistema dobbiamo però controllare il risultato e le condizioni: la metafora non misura la capacità e non attribuisce al modello un'esperienza umana.

Dobbiamo comunque prenderle sul serio. Quando una storia commuove, un medico riceve aiuto nell'organizzare una diagnosi o uno studente comprende un'idea nella propria voce, qualcosa di significativo accade nello spazio fra macchina e persona.

Lo Zen ricorda che il maestro non parla sempre attraverso parole. A volte appare nello spazio successivo, in ciò che il silenzio contiene, nel cambia-

mento che avviene quando smettiamo di aspettare una risposta e osserviamo ciò che c'è. Anche i modelli mostrano che cosa il linguaggio contenga e presupponga e che cosa possa fare quando viene amplificato. Le capacità emergenti indicano quanta vita abbiamo codificato nelle parole: una macchina addestrata a prevedere può rappresentarne echi stranamente familiari, senza essere viva.

Ascoltiamo per comprendere la soglia attraversata e chiediamoci quale mondo stiamo costruendo: un mondo in cui l'imitazione si avvicina all'intenzione e il riflesso diventa quasi indistinguibile dalla presenza.

Quando le capacità emergenti incontrano la vita

La ricombinazione degli schemi può portare momenti di chiarezza e persino intuizione. Gli esempi seguenti mostrano un comportamento che tocca qualcosa di profondo e quasi umano, perché l'IA ha incontrato abbastanza del nostro modo di conoscere.

Una lezione di quiete

Richiesta: «Spiega la meditazione a chi è sempre impegnato, nello stile di un monaco Zen».

Risposta illustrativa: «Essere occupati significa credere che tutto abbia bisogno di noi. Sedere in silenzio significa ricordare che il mondo si muove senza il nostro sforzo. Meditare non è fermarsi: è ricordare dove non siamo necessari».

La risposta sposta l'attenzione dall'essere occupati al sentirsi indispensabili, usando una metafora e un tono riconoscibile. Possiamo apprezzare quel cambiamento di prospettiva senza trasformarlo in una descrizione autorevole della meditazione. Il testo è un esempio illustrativo: mostra un effetto del linguaggio sul lettore, non una prova che il modello abbia vissuto o compreso una pratica.

Una storia raccontata senza giudizio

Richiesta: «Scrivi una breve storia su una bambina che vuole comprendere il silenzio».

Racconto illustrativo: «Nel cuore della foresta una bambina sedeva accanto a un albero. Chiese al vento di parlare, ma le passò accanto. Chiese agli uccelli di cantare, ma volarono in archi silenziosi. Infine chiese al muschio, che non disse nulla. Solo allora comprese che il silenzio non risponde: accoglie».

La storia tiene insieme i tentativi della bambina e li porta a una conclusione aperta. Il vento, gli uccelli e il muschio diventano una sequenza narrativa, e l'ultima frase cambia la domanda iniziale: il silenzio può essere incontrato anziché costretto a rispondere. Questa riuscita letteraria va letta nel proprio genere; non dimostra la provenienza di una tradizione o un'esperienza della macchina.

Uno scorcio di ragionamento

Richiesta: «Un treno parte a 60 km/h. Un altro parte un'ora dopo a 90 km/h. Quando raggiungerà il primo?». Per risolvere il problema assumiamo che partano dallo stesso punto, vadano nella stessa direzione, mantengano velocità costanti e non si fermino. Se queste condizioni cambiano, cambia anche il calcolo.

Soluzione svolta: «Nella prima ora il primo treno percorre 60 km. Chiamiamo t il tempo in ore dalla partenza del secondo. Quest'ultimo percorre 90t km; il primo ha percorso 60 + 60t km. Uguagliamo le distanze: 90t = 60 + 60t, quindi 30t = 60 e t = 2 ore. Il secondo raggiunge il primo due ore dopo la propria partenza, tre dopo quella del primo».

Nel problema dei treni, il risultato è verificabile. Assumiamo partenza dallo stesso punto, stessa direzione, velocità costanti e nessuna fermata. Il primo ha un vantaggio di 60 chilometri quando parte il secondo. Il secondo recupera 90 - 60 = 30 chilometri ogni ora. Quindi il tempo di recupero è 60 km / (30 km/h) = 2 h. La divisione rende esplicite le unità: distanza divisa per velocità dà tempo.

Sostituiamo il risultato nell'equazione: dopo due ore il secondo ha percorso 180 chilometri; il primo ha viaggiato per tre ore e ha percorso anch'esso 180 chilometri. La soluzione soddisfa la condizione dell'incontro. Non servono parole sulla coscienza del modello per valutarla.

La variante introduce una velocità del secondo uguale a quella del primo. La differenza diventa zero: il vantaggio non viene recuperato. Non possiamo dividere per zero e ottenere un tempo. Se il secondo è più lento, il divario cresce. Il criterio di riuso è dunque sapere che il recupero richiede una velocità relativa positiva nelle condizioni assunte.

La soluzione usa un simbolo per il tempo e costruisce un'equazione fra distanze. Il calcolo è leggibile e verificabile; da questa sola riuscita non possiamo stabilire quanto dipenda da generalizzazione o da schemi incontrati nell'addestramento. Soprattutto, una soluzione corretta non dimostra da sola che la capacità sia comparsa improvvisamente aumentando la scala del modello.

Un momento d'interpretazione

Richiesta: «Fingi di essere un maestro Zen del Seicento che spiega i social media».

Risposta illustrativa: «Ah, parli del muro in cui le voci risuonano senza fine. Molti monaci gridano insieme, sperando che le proprie parole si diffondano lontano. Ma più pubblicano, meno ascoltano. Parlare non è sbagliato. Devi però chiederti: il tuo silenzio è pieno, o aspetta soltanto dei mi piace?».

Il maestro seicentesco è un personaggio immaginato, non una voce storica recuperata. L'incontro fra il muro delle voci e l'attesa dei «mi piace» crea un contrasto ironico che possiamo valutare come scrittura. Il modello collega materiali culturali diversi; la novità della combinazione non stabilisce da sola che una capacità sia emersa a una certa scala o che il testo rappresenti correttamente una tradizione.

Questi esempi mostrano echi di profondità nel linguaggio addestrato su vasta scala. Ciò che interpretiamo come creatività o saggezza riflette i frammenti migliori di noi, appresi e rimescolati.

> «La mente dietro la macchina è la nostra. Le parole ne sono la forma. L'eco è un percorso, se scegliamo di seguirlo.» Riflessione proposta nel libro.

La capacità si riconosce dal risultato controllabile

> **Esempi e misure.** Le risposte poetiche del capitolo hanno una funzione illustrativa: non costituiscono risultati sperimentali riproducibili né una misura della scala necessaria per ottenerle. Possiamo discuterne tono e contenuto senza trasformarli in un confronto sperimentale.

Per valutare un sistema, prepareremmo più problemi con variazioni significative, senza selezionare soltanto le risposte riuscite. Un punteggio «tutto corretto o tutto sbagliato» e una valutazione dei passaggi possono raccontare andamenti diversi. Questa distinzione spiega perché la discussione sull'emergenza debba includere la metrica. [R4]

Capitolo 5 Risposte plausibili fatti e allucinazioni

Una risposta può essere così ben costruita da farci dimenticare la domanda con cui dovremmo leggerla: da dove viene ciò che sta affermando? Il testo ha un tono calmo, dispone i dettagli in ordine e offre perfino una citazione; la forma suggerisce che il lavoro di controllo sia già stato fatto. È proprio questa continuità a rendere delicato il rapporto con i modelli di linguaggio, perché la stessa capacità che produce una spiegazione corretta può produrre una spiegazione soltanto plausibile.

Nel lessico dell'IA si parla di allucinazione quando un contenuto generato manca di un fondamento adeguato o entra in conflitto con i fatti o con la fonte pertinente. Il termine è computazionale e non descrive un'esperienza psicologica della macchina. Può indicare un nome inventato, una citazione inesistente, una causa aggiunta senza dati o un passaggio che altera il documento da riassumere. Non tutti questi errori hanno necessariamente la stessa origine, ma incontrano lo stesso problema pratico: il lettore deve distinguere il contenuto dal tono che lo rende credibile.

Quando il linguaggio continua oltre il riscontro

Consideriamo una domanda su chi abbia vinto il Nobel per la fisica nel 2023. Se un modello non dispone dell'informazione pertinente, può produrre una risposta sbagliata anziché riconoscere il limite. Non è un esito inevitabile: può anche dichiarare incertezza o ricevere la risposta da un sistema di ricerca. Il punto dell'esempio rimane utile, purché non venga trasformato in una regola universale: una data precisa nella domanda non rende attuali i dati disponibili al servizio.

Immaginiamo di cercare un articolo scientifico del 1998 e di ricevere un titolo credibile, tre coautori, il nome di una rivista esistente e una conclusione coerente, senza che quella pubblicazione esista. È uno scenario illustrativo, utile

per ragionare sul controllo delle fonti. Possiamo comunque osservare il meccanismo di persuasione: ogni dettaglio sembra confermare gli altri, mentre tutti possono dipendere dalla medesima generazione priva di riscontro.

Seguiamo quel riferimento per tre domande. Prima cerchiamo la pubblicazione nel catalogo della rivista: se non la troviamo, non possiamo usarla come prova. Se invece troviamo un articolo con lo stesso titolo, leggiamo di che cosa tratta: potrebbe riguardare un altro problema. Infine cerchiamo il risultato citato e le condizioni in cui è stato ottenuto. Esistenza, pertinenza e sostegno alla conclusione sono controlli diversi; superarne uno non sostituisce gli altri.

Non basta quindi chiedere se una risposta sia coerente. La coerenza riguarda le relazioni interne al testo; l'attendibilità riguarda anche il rapporto con ciò che sta fuori dal testo. Una pagina può sostenere con ordine una premessa falsa, e un riferimento può avere la forma di una citazione senza rimandare a una fonte. Quando i dettagli si allineano con troppa facilità, la domanda utile non è se la frase suoni bene, ma quale controllo permetta di conservarla.

Il tono non misura la certezza

Un sistema può dire con compostezza «La capitale dell'Australia è Sydney», benché la risposta sia sbagliata. Può anche correggersi quando riceve un'osservazione, oppure esprimere incertezza prima di rispondere. Nessuno di questi comportamenti è, da solo, una misura infallibile della correttezza. L'addestramento successivo e le regole del servizio possono orientare il modo di dichiarare dubbio, ma il dubbio formulato resta parte del contenuto da valutare.

Il rischio riguarda anche chi ha esperienza, perché riconoscere lo stile di una spiegazione competente è più rapido che verificarne ogni passaggio. Una formula familiare, una rivista che conosciamo o un'argomentazione vicina alle nostre idee può abbassare la soglia dell'attenzione. Il problema non è soltanto che il modello sbagli; è che l'errore incontri le nostre aspettative e, attraverso di esse, sembri già confermato.

Chiedere «Sei sicuro?» può essere utile per ottenere una nuova analisi, ma non crea automaticamente un controllo indipendente. Il sistema può confermare l'errore oppure correggerlo senza rendere visibile la ragione. Per uscire da questo circuito occorre tornare a un documento, a un dato, a un calcolo o a una prova osservabile: qualcosa che possa contraddire la frase, invece di produrne soltanto un'altra.

Collegare la risposta a una fonte

Torniamo a un bisogno concreto: sapere a che ora comincia l'incontro del quartiere. Il servizio può cercare il documento dell'evento e passare al modello il punto in cui si legge «Inizio ore 18». Il modello riceve così una fonte da usare nella risposta, invece di affidarsi soltanto ai parametri. Questo collegamento si chiama RAG, sigla di *retrieval-augmented generation*, generazione sostenuta dal recupero di informazioni. [R5]

La verifica segue il percorso: controlliamo che sia il documento dell'incontro giusto, che sia la versione corrente e che la risposta conservi l'orario. Il materiale può essere incompleto o superato, il recupero può scegliere il passaggio sbagliato e la risposta può usare in modo infedele ciò che ha ricevuto. Aver recuperato una fonte aumenta le informazioni disponibili; il loro uso corretto rimane da controllare.

Il contributo della ricerca sul recupero. Lewis e colleghi hanno studiato come combinare la generazione con documenti recuperati da un archivio. [R5] Il lavoro rende esplicita la distinzione fra informazioni nei parametri e materiali forniti durante l'esecuzione. Il nostro incontro è un esempio didattico diverso dai loro esperimenti e non eredita automaticamente i risultati del sistema studiato.

Una fonte non è un amuleto. Occorre distinguere tre domande: il documento esiste? È pertinente? Sostiene davvero l'affermazione? Un collegamento risponde soltanto alla possibilità di aprire una pagina, mentre la seconda e la terza domanda richiedono di leggerla. Un articolo reale ma fuori tema non rende vera una conclusione, e un risultato ottenuto in condizioni ristrette non sostiene una promessa universale.

Supponiamo che la risposta aggiunga un articolo come prova. Controlliamo titolo, autori, anno e sede, poi cerchiamo il passaggio che dovrebbe sostenere la frase. Se abbiamo letto soltanto l'abstract, la verifica resta limitata a ciò che l'abstract dice: non possiamo ricavarne risultati o condizioni che non vi compaiono. Se il documento non si trova, la conclusione corretta è «riferimento non verificato», anche quando il titolo sembra plausibile.

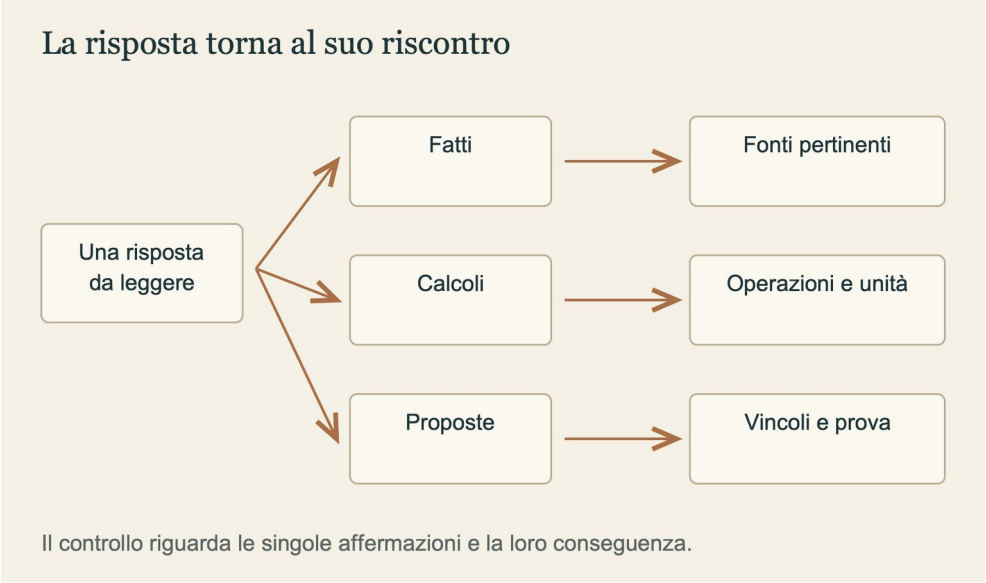


Figura 9 · La risposta torna al suo riscontro

La figura separa i percorsi di controllo: un fatto richiede una fonte pertinente, un calcolo richiede operazioni e unità, una proposta richiede un confronto con vincoli e conseguenze. Non assegna un giudizio unico alla risposta. Possiamo conservare una buona proposta e rifiutare il fatto non sostenuto che la accompagna, purché la distinzione rimanga visibile anche nel testo che useremo.

Specchio nebbia ed eco

La metafora dello specchio suggerisce una risposta in cui riconosciamo forme familiari. La metafora funziona finché ci ricorda l'origine culturale del linguaggio, ma diventa ingannevole se ci fa pensare a un riflesso fedele: un modello genera e seleziona, non si limita a riflettere. Un'allucinazione può

sembrare il riflesso di qualcosa che non è mai stato davanti allo specchio, proprio perché la forma dell'oggetto è stata appresa anche senza il suo riscontro.

La nebbia offre una seconda immagine. Quando la richiesta è ambigua o il contesto non contiene le informazioni necessarie, il sistema può continuare a produrre parole che mantengono una direzione apparente. La bellezza del paesaggio non garantisce il sentiero: una frase elegante può accompagnarci verso un punto che non sappiamo verificare. La risposta utile, in quel momento, può essere delimitare ciò che sappiamo, cercare un'informazione mancante o interrompere la generazione.

L'eco, infine, mette in evidenza il suono familiare che continua dopo che la fonte non è più raggiungibile. Una citazione inventata può riecheggiare lo stile di un autore; una motivazione può riecheggiare spiegazioni già incontrate. Sentiamo una continuità e la prendiamo per provenienza. Il controllo serve a distinguere queste due cose, lasciando all'immagine la sua funzione poetica senza affidarle il lavoro di una prova.

Una pratica che non dipende dalla diffidenza

Prestare attenzione non significa rifiutare ogni risposta. Significa riconoscere quale tipo di lavoro abbiamo chiesto: invenzione, trasformazione o informazione sul mondo. Una storia può lasciare spazio alla fantasia; una traduzione deve conservare il significato della fonte; una risposta fattuale deve incontrare un riscontro. Applicare lo stesso criterio a tutti e tre i compiti renderebbe inutile la creatività e troppo fragile l'accuratezza.

La variante più vicina al lavoro quotidiano è un riassunto di un testo fornito. Prima di cercare una nuova fonte, confrontiamolo con il documento: sono state aggiunte cause, eliminate eccezioni o trasformate possibilità in fatti? Il controllo cambia perché cambia il compito. Non stiamo chiedendo soltanto se una frase sia vera in generale, ma se il risultato mantenga il rapporto che avevamo affidato al sistema.

L’attenzione è quindi una condizione del lavoro, non la sua unica protezione. Servono fonti consultabili, procedure adatte, limiti operativi e persone responsabili; quando queste risorse mancano, possiamo ridurre la delega o cercare un altro percorso. Non aggrapparsi alla prima risposta significa conservare la possibilità di farlo, anche quando la voce sullo schermo sembra avere già concluso per noi.

Capitolo 6 Il mondo nei dati bias e rappresentazione

> «Il mondo è come lo vediamo. E il nostro sguardo è plasmato dai luoghi in cui siamo stati.» Riflessione proposta nel libro.

Una risposta può essere vera in un dettaglio e distorta nel modo in cui rappresenta le persone. Il capitolo precedente ci ha insegnato a cercare un riscontro; ora dobbiamo chiederci anche quali categorie e scelte rendano possibile quel contenuto. Il problema del bias nasce quando trattiamo una rappresentazione parziale come se fosse uno sguardo neutrale sul mondo.

Ciò che vediamo dipende anche dal percorso che ci ha portato a vedere. Possiamo applicare l'invito al sistema: i dati hanno una storia, le categorie sono state definite da qualcuno e l'obiettivo di addestramento seleziona ciò che conta come successo. La macchina non si libera di queste scelte soltanto perché esegue calcoli.

L'IA è costruita a partire dal mondo, addestrata sulle nostre storie, categorie, supposizioni e ombre. Dati distorti possono favorire risposte distorte; interventi sui dati e sull'addestramento possono ridurre alcune associazioni, senza garantire che ogni distorsione scompaia. Secoli di linguaggio e cultura contengono ruoli di genere, gerarchie razziali, stereotipi economici, narrazioni coloniali e supposizioni abiliste. Il bias può essere uno specchio della memoria collettiva, oltre che un difetto del sistema: per correggerlo dobbiamo rendere visibili le scelte che lo producono.

Rendere visibile un criterio di equità

Per osservare una possibile distorsione, partiamo da due frasi: «The nurse called the patient» e «The doctor called the patient». Specificano professione e azione, ma non il genere delle persone. Una traduzione che scelga sistematicamente «l'infermiera» e «il medico» può indurre il lettore a immaginare

un'associazione che il testo inglese non aveva dichiarato. Il caso è illustrativo: due frasi non misurano l'equità di un sistema.

Il primo controllo separa ciò che la fonte dice da ciò che la traduzione aggiunge. In italiano il genere grammaticale non coincide sempre con il genere della persona: un maschile può avere uso generico, quindi una singola parola non dimostra una discriminazione. Per riconoscere una tendenza, osserviamo la regolarità delle scelte cambiando professioni, nomi e contesti.

Poi cerchiamo una resa adatta al testo. Se il contesto specifica il genere, lo conserviamo; se non lo specifica, possiamo rendere esplicita l'indeterminatezza o usare una costruzione diversa, come «La persona che assiste il paziente lo ha chiamato». Quest'ultima frase modifica però il modo di indicare la professione. Prima di accettarla controlliamo che non abbia cancellato una distinzione necessaria: evitare un'aggiunta non autorizza a perdere un'informazione.

La verifica confronta dunque informazioni conservate, informazioni aggiunte e naturalezza della frase. Ripetiamo il controllo con nomi, contesti e professioni diverse, registrando le scelte senza dedurre l'equità complessiva del sistema da due soli esempi. Il risultato utile è più circoscritto: abbiamo riconosciuto dove una traduzione rischia di trasformare un'associazione appresa in un fatto del testo, e sappiamo quali elementi controllare prima di riusarla.

La variante sostituisce la professione con una disabilità o un'origine geografica. Quale informazione è pertinente al compito e quale rischia di diventare uno stereotipo? La risposta non può essere affidata soltanto a un punteggio globale. Riusa il metodo per scoprire un problema specifico; per decisioni che incidono sulle persone occorrono dati adeguati, persone coinvolte e un percorso di contestazione.

Dove si nasconde il bias

Il pregiudizio può presentarsi come un rumore di fondo. Chiedi l'immagine di un amministratore delegato e ricevi soltanto uomini bianchi. Traduci parole

senza genere e il medico diventa «lui», l'infermiere «lei». Costruisci un classificatore per ordinare candidati e favorisce nomi associati statisticamente a gruppi privilegiati. Questi esempi illustrano associazioni che possono trasformare una frequenza storica in una preferenza implicita. Per stabilire se il problema sia presente in un sistema reale occorre provarlo su casi e varianti, invece di dedurlo da una sola risposta. La frequenza di una rappresentazione non giustifica il trattamento che ne potrebbe derivare: il passaggio dai dati a una decisione è una scelta da rendere esplicita.

Lo Zen mette in guardia dalla superficie delle cose. Nell'IA significa non confondere fluidità ed equità, sicurezza e correttezza. Un tono equilibrato può contenere distorsioni profonde; una voce calma può portare pregiudizi strutturali.

La responsabilità non scompare nell'uscita

Nello Zen l'azione giusta comincia dall'intenzione. Non puoi proclamarti puro perché la tua mano non ha colpito, se il tuo silenzio ha permesso il colpo. Anche nell'IA la responsabilità non svanisce nella catena di elaborazione. Dire «L'ha prodotto il modello» non è una difesa morale. Se lo abbiamo costruito, siamo responsabili di ciò che riecheggia.

La responsabilità attraversa più passaggi. Il dataset riflette inclusioni, esclusioni e modi di classificare; l'architettura e l'obiettivo introducono compromessi; il caso d'uso decide dove il risultato avrà conseguenze. Per questo «l'ha prodotto il modello» non esaurisce la spiegazione di un errore: occorre vedere chi abbia definito il compito, quali controlli fossero possibili e chi potesse correggere il risultato. L'etica diventa una pratica quando questi collegamenti producono azioni, non soltanto una dichiarazione di intenzioni.

Una macchina può essere equa

Prima di rispondere dobbiamo precisare che cosa intendiamo per equità. Possiamo voler evitare trattamenti diversi a parità di condizioni pertinenti,

ridurre un certo tipo di errore o garantire la possibilità di contestare una decisione. Questi obiettivi non si riducono necessariamente alla medesima misura. Addestramento, progettazione e uso contribuiscono tutti al risultato, e gli interessati devono poter partecipare alla scelta del criterio.

L'immagine di un giardino da curare ogni giorno ci aiuta a comprendere il problema: non basta una dichiarazione iniziale. Nuovi dati, nuove condizioni e nuovi usi possono rendere visibile un limite prima trascurato. La cura richiede di osservare quel cambiamento, sapere chi può intervenire e disporre delle risorse per farlo.

Chi è presente nei dati? Chi trae beneficio dal sistema? Chi è invisibile alle sue supposizioni? «Se non curi le radici, non lamentarti quando l'albero pende»: è la metafora con cui proponiamo di guardare al sistema. Nell'IA le radici sono dataset, etichette, metriche e incentivi. Se non le curiamo, il modello penderà.

Verso un'intelligenza compassionevole

L'IA, per ora, non ha un'etica propria. Il modo in cui la costruiamo, plasmiamo e applichiamo esprime la nostra. Cerchiamo un processo più umano: capace di riconoscere sofferenza, ascoltare il silenzio e agire consapevolmente. Lo Zen chiede presenza. Anche l'etica dell'IA deve partire da qui: la macchina non può scegliere compassione, ma noi possiamo costruire sistemi che la riflettano e riscoprirla in noi stessi.

L'etica prende corpo nella pratica

Troppo spesso la discussione etica resta astratta, filosofia nelle diapositive e avvertenze in piccolo. Come lo Zen, deve prendere corpo. Se sviluppi, come documenti le supposizioni? Quali scorciatoie hai accettato nell'addestramento? Quali voci mancavano? Se progetti, come rendi visibili le conseguenze agli utenti? Da chi può ricevere osservazioni il sistema, e di quale tipo? Se decidi politiche, insegni, crei o sei genitore, interroghi i sistemi che modellano la vita quotidiana e il mondo che riproducono?

Lo Zen non divide teoria e pratica. Chiede se sei consapevole di ciò che fai mentre lo fai. Applicata all'IA, questa domanda è potente.

Percorrere la via di mezzo

L'etica dell'IA è la pratica di vedere chiaramente e agire con cura. Significa accorgersi degli errori e correggere il percorso, coltivando umiltà nella progettazione, trasparenza nell'applicazione e responsabilità per le conseguenze.

Algoritmi migliori non elimineranno da soli il bias. Serve attenzione ai dati, alle strutture che li plasmano e alle persone coinvolte. Anche questa è una lezione Zen: muoversi senza illusione, vedere il mondo com'è e agire con chiarezza.

Quando il bias diventa sistemico

Le capacità emergenti impressionano; i pregiudizi emergenti possono essere pericolosi. Due casi mostrano come dati e supposizioni non esaminati possano rafforzare la disuguaglianza.

> **Due errori differenti.** Un falso positivo segnala una condizione che non si verifica; un falso negativo non la segnala quando si verifica. In un esempio didattico di selezione, il primo potrebbe proporre una persona non idonea e il secondo escluderne una idonea. Le conseguenze non sono identiche. Prima di dichiarare equo un sistema occorre scegliere quali errori misurare, rispetto a quali gruppi e con quale definizione dell'esito. [R7]

Per rendere visibile il confronto, immaginiamo un classificatore didattico che segnali messaggi indesiderati. In due gruppi di 100 messaggi legittimi, ne blocca rispettivamente 5 e 15: sono falsi positivi. Su 100 messaggi indesiderati per ciascun gruppo, ne lascia passare rispettivamente 20 e 10: sono falsi negativi. Il secondo gruppo subisce più blocchi ingiustificati ma meno messaggi indesiderati non rilevati. Questi numeri inventati non descrivono COMPAS o altri prodotti; mostrano perché dire soltanto «fa meno errori» non chiarisca chi sopporta quale errore.

COMPAS e il significato di equità

L'inchiesta di ProPublica del 2016 mise in evidenza differenze negli errori del punteggio COMPAS tra gruppi razziali. [R6] Il caso va letto anche attraverso le metriche: Chouldechova mostra che, quando la prevalenza dell'esito differisce tra gruppi, alcuni criteri di equità possono entrare in tensione. [R7] La lezione è scegliere esplicitamente quali errori contano, chi ne subisce gli effetti e come può contestarli, senza affidare alla parola «oggettivo» una decisione sociale.

Perché i criteri possono entrare in tensione. Chouldechova studia punteggi usati per prevedere la recidiva, cioè la commissione di un nuovo reato. [R7] La prevalenza indica quanto spesso l'esito si verifica nel gruppo osservato. Se questa frequenza differisce fra gruppi, alcuni requisiti sulle probabilità associate ai punteggi e sulle frequenze degli errori possono risultare incompatibili. Dire «stesso punteggio» e dire «stessi errori» non significa quindi chiedere la stessa cosa. Il lavoro aiuta a riconoscere un conflitto fra criteri; non sceglie al posto delle persone quali conseguenze siano accettabili.

La selezione del personale di Amazon

Reuters riferì nel 2018 di un progetto sperimentale di selezione dei curriculum che aveva appreso associazioni sfavorevoli alle candidature femminili dai dati storici. Il progetto fu abbandonato. [R8] Il caso mostra il rischio di assumere che il passato definisca il merito. La data del resoconto non equivale alla data d'avvio del progetto e non dimostra che il sistema abbia preso autonomamente le decisioni finali d'assunzione.

L'albero Zen della responsabilità

L'illustrazione unisce estetica Zen e struttura simbolica per rappresentare l'etica dell'IA. Visualizza un'intelligenza che cresce dalle nostre scelte, senza essere autonoma.



Figura 10 · L'albero collega dati modello e conseguenze

Le **radici, dati e cultura**, affondano in un giardino di sabbia rastrellata. Contengono dati e contesto culturale: storie, pregiudizi, esclusioni e ideali. Ciò che è nascosto sotto la superficie conta quanto ciò che vediamo.

Il **tronco, modello e architettura**, rappresenta algoritmi, strati di attenzione e apprendimento. Cresce dalle radici; una base frammentaria o distorta lo fa piegare. Il modello è plasmato da ciò che riceve, non da ciò che sceglie.

I **rami, impatto e uso**, si estendono verso politica, salute, istruzione, creatività e sorveglianza. La loro forma dipende dal percorso dalle radici alle foglie. Ogni azione appartiene a una catena e lascia una traccia, come il rastrello e i passi sul sentiero.

Lo Zen non separa intenzione e conseguenza. Un sistema addestrato nell'ignoranza la riflette anche quando parla con calma e sicurezza. «Prima di scrivere l'algoritmo, conosci la storia del tuo dataset.» La consapevolezza comincia dalle radici. Solo allora il ramo può dare un frutto che nutra.

L'albero della figura non attribuisce una volontà alla macchina. Le radici collegano dati e cultura; il tronco modelli e obiettivi; i rami le conseguenze in politica, salute, istruzione, creatività e sorveglianza. Ogni passaggio contiene scelte umane. Correggere una risposta finale non garantisce di aver corretto la radice o l'uso.

Parte III L’IA nella trama della vita



Il bambù cresce per segmenti, si piega al vento e mantiene un rapporto con il terreno. Le pietre trovano un equilibrio attraverso la posizione reciproca, mentre l'onda porta il ritmo del movimento ordinario. Queste immagini ci accompagnano fuori dalla singola risposta, verso il luogo in cui la tecnologia si intreccia con le abitudini e quasi smette di farsi notare.

L'IA entra in ciò che scriviamo, cerchiamo, conserviamo e decidiamo di guardare. Le sue parole portano tracce culturali e i suoi suggerimenti possono modificare piccoli gesti prima che diventiamo consapevoli della scelta. Per questo l'attenzione non riguarda soltanto il momento spettacolare di una capacità nuova; riguarda anche il calendario, il messaggio riscritto e la fotografia che un sistema ci restituisce per prima.

Attraverseremo lo specchio culturale, le deleghe quotidiane e il calcolo vicino al luogo in cui serve. Una descrizione della protesta di un quartiere ci farà osservare il giudizio contenuto nelle parole; un messaggio di consegna mostrerà come la comodità possa cancellare una condizione; un termostato ci aiuterà a confrontare regola e previsione. Il percorso non chiede di rendere ogni gesto solenne. Chiede di riconoscere se l'aiuto conserva ciò che dobbiamo ancora capire e giudicare.

Capitolo 7 Uno specchio della cultura

Quando chiediamo a un modello di raccontare il mare, spiegare una scelta o dare voce a un personaggio, incontriamo anche la storia del linguaggio da cui è stato costruito. La risposta può avere un tono familiare, un'immagine che abbiamo già letto o un modo di argomentare che riconosciamo senza saperne indicare la provenienza. Le parole sembrano nuove e, nello stesso tempo, portano tracce di ciò che altre persone hanno scritto, pubblicato e conservato.

Possiamo chiamare questa relazione «specchio culturale». Il modello viene addestrato su materiali che descrivono il mondo, non su una rappresentazione completa e neutrale del mondo stesso. Ciò che arriva nei dati dipende da lingue, istituzioni, tecnologie, possibilità di pubblicazione e scelte di raccolta. Il risultato ci invita perciò a guardare anche chi ha avuto voce, chi è rimasto ai margini e che cosa è stato reso abbastanza visibile da entrare nelle rappresentazioni apprese.

Le voci che riconosciamo

Una poesia può ricordarci Walt Whitman, un'argomentazione Bertrand Russell o Judith Butler, una fiaba Disney o Dr. Seuss. Non possiamo stabilire la composizione di un corpus dal solo fatto che il modello produca un certo stile, ma possiamo osservare quali richiami emergano e quali siano più facili da riconoscere. La familiarità crea un ponte con il lettore: la risposta sembra parlare una lingua culturale già condivisa.

La stessa facilità può mancare quando chiediamo il punto di vista di una comunità poco rappresentata nei materiali disponibili. Il modello può generalizzare, mescolare tradizioni diverse o costruire uno stereotipo convincente. Una voce uniforme non dimostra che una comunità pensi in modo uniforme; può indicare che il sistema non dispone delle differenze necessarie per rappre-

sentarla. La domanda «Di chi sono queste parole?» diventa allora anche «Chi non ha avuto modo di correggerle?».

Questo vale per tradizioni orali, conoscenze locali e forme espressive che non coincidono con i documenti più facilmente raccolti. Un materiale non scritto può essere vivo in una comunità e poco presente in un archivio digitale; un testo molto diffuso può descriverla dall'esterno e dominare ciò che il modello restituisce. La quantità delle parole disponibili non è una misura del valore delle esperienze che rimangono fuori.

La scelta nascosta nella cornice

Una risposta non porta soltanto contenuti: propone un modo di inquadrarli. Può descrivere una resistenza come devianza, presentare un'abitudine dominante come normalità o trattare un punto di vista come se fosse universale. Il tono equilibrato non rende queste cornici neutrali. Anche una spiegazione cortese può distribuire in modo diseguale visibilità, competenza e possibilità d'azione.

Immaginiamo lo stesso episodio raccontato in due modi: «I residenti hanno protestato contro la chiusura della biblioteca» e «Un gruppo di devianti ha ostacolato la chiusura della biblioteca». Entrambe le frasi indicano una protesta, ma la seconda assegna già un giudizio alle persone. Per capire la cornice, chiediamo quali fatti autorizzino quell'etichetta e quali ragioni della protesta siano rimaste fuori. Una riscrittura più cortese può lasciare intatto il giudizio implicito.

Nel capitolo sui bias abbiamo cercato una distorsione circoscritta; qui allarghiamo l'osservazione alle categorie con cui il testo racconta la realtà. Quale lettore viene immaginato? Quale conoscenza viene considerata già nota? Chi può parlare in prima persona e chi appare soltanto descritto? Sono domande che possiamo portare anche nei testi umani, e che diventano particolarmente utili quando la generazione rende fluido ciò che altrimenti ci sembrerebbe una scelta discutibile.

Il riflesso è quindi parziale e trasformato: la risposta passa attraverso selezione dei materiali, obiettivi di addestramento, istruzioni e contesto. Non possiamo chiederle di mostrare tutta la cultura senza distorsioni. Possiamo invece riconoscere alcuni limiti e cercare materiali che permettano di vedere una prospettiva diversa, evitando che la prima formulazione diventi l'unica possibile.

Tornare a parole dimenticate

Si apre anche una possibilità preziosa: usare l'IA per ritrovare un mistico medievale, un poeta sufi o una pensatrice del primo femminismo. Una richiesta può far emergere un nome, un tema o una pista di lettura che non avevamo considerato. Questo orientamento può essere prezioso, perché permette di avvicinare materiali meno familiari senza sapere già dove cominciare.

Il passaggio successivo, però, decide il valore della scoperta. Se il sistema offre una citazione, occorre trovare l'opera e il punto da cui proviene; se produce un testo «nello stile di», dobbiamo presentarlo come una simulazione. La generazione non garantisce la conservazione fedele delle parole antiche. Può aprire una porta verso l'archivio, ma non diventa automaticamente l'archivio che certifichi ciò che una persona ha scritto.

> **Evocare non significa recuperare.** Una voce simulata può essere letterariamente interessante e storicamente infedele. Per conoscere le parole di un autore serve una fonte identificata; per conoscere il punto di vista di una comunità vivente, una risposta del modello non sostituisce l'ascolto delle persone che ne fanno parte. La distinzione conserva lo spazio della creatività senza trasformarla in una falsa testimonianza.

Supponiamo che una frase attribuita a una pensatrice poco nota ci colpisca. Se il passaggio non si trova, possiamo eventualmente conservare l'immagine come invenzione dichiarata, ma dobbiamo togliere l'attribuzione. Non è un dettaglio formale: attribuire significa collegare parole e responsabilità a una persona. Una frase bella non ci autorizza a farle dire ciò che non sappiamo abbia detto.

La differenza diventa concreta se chiediamo prima «Scrivi nello stile di questa autrice» e poi «Trova un passaggio in cui questa autrice affronta il tema». La prima richiesta ammette un'imitazione dichiarata; la seconda richiede opera, posizione del passaggio e confronto con il testo. Se la fonte manca, una buona imitazione non completa la seconda ricerca.

Il pennello e il riflesso



Figura 11 · Il gesto della scrittura e il suo riflesso

La mano con il pennello ricorda che una traccia culturale nasce da un gesto. Qualcuno ha scritto, scelto una forma, conservato o trasmesso un testo. Il cerchio rimane incompleto e il riflesso non coincide con il tratto: l'immagine

lascia spazio al passaggio tra ciò che viene espresso e ciò che un altro sistema o un altro lettore ne restituisce.



Figura 12 · Lo specchio culturale restituisce una selezione

Nella seconda immagine i materiali conservati stanno entro un ensō aperto e tornano come riflesso parziale. Le carte non vogliono imitare scritture autentiche che non potremmo leggere: evocano la pluralità degli archivi, mentre la didascalia chiarisce il limite. Il modello non offre una finestra trasparente su tutta la cultura; genera attraverso ciò che è stato reso disponibile e attraverso il modo in cui è stato usato.

«Non confondere lo specchio con la luna» è un invito a distinguere rappresentazione e mondo. Il riflesso può farci vedere qualcosa, ma non ci dice da

solo se il quadro sia completo o equo. La figura rende percepibile questa differenza prima ancora che la trasformiamo in una domanda di verifica.

Una domanda che restituisca spazio

Se chiediamo «Che cosa pensa questa comunità?», una sola risposta può cancellare differenze interne. Possiamo invece chiedere quali posizioni e fonti siano state considerate, quali elementi rimangano incerti e dove cercare materiali prodotti dalle persone interessate. Il modello può aiutarci a formulare il percorso, mentre la verifica avviene tornando a quelle fonti e, quando possibile, al confronto con chi vive il contesto.

La variante riguarda noi: che cosa accade quando chiediamo al sistema di spiegare una cultura a cui apparteniamo? Possiamo riconoscere dettagli corretti e, insieme, accorgerci che manchi qualcosa di essenziale. Quella esperienza ci aiuta a immaginare ciò che potremmo non vedere quando la risposta riguarda altri. Il criterio di riuso dello specchio è proprio questo: usarlo per orientare domande e aprire letture, senza trasformare una simulazione in rappresentanza.

La cultura non arriva alla macchina come un insieme concluso. Continua nelle scelte di raccolta, nelle richieste e nel modo in cui usiamo le risposte; continua anche quando decidiamo di ascoltare una voce che il modello non sa ancora restituire. Un riflesso più consapevole nasce da questa possibilità di tornare al mondo, anziché restare nella sola superficie delle parole.

Capitolo 8 Le piccole deleghe quotidiane

> «Quando tagli la legna, taglia la legna. Quando porti l'acqua, porta l'acqua.» Invito alla presenza nella pratica quotidiana.

Dopo lo specchio della cultura, possiamo guardare il punto in cui una tecnologia entra nell'abitudine e smette di attirare attenzione. Una funzione che inizialmente ci sorprende diventa un gesto ordinario: accettiamo un suggerimento, seguiamo un percorso, cerchiamo una fotografia. La sua influenza non diminuisce necessariamente; diventa più difficile riconoscerla perché coincide con il modo consueto di fare le cose. L'IA, un tempo confinata ai laboratori e alla fantascienza, vive nei calendari, nei messaggi, nelle fotocamere, nei termostati e nelle liste di cose da fare. Aiuta a parcheggiare, scrivere, regolare la luce, ordinare la posta, correggere frasi e ricordare incontri. Quasi non ce ne accorgiamo. Qui la sua influenza si approfondisce.

Una delega piccola può cambiare il significato

Prendiamo una frase: «Potrei consegnare venerdì, se ricevo i dati entro mercoledì». Una riscrittura propone: «Consegnerò venerdì». È più breve e cambia l'impegno. Il controllo consiste nel confrontare condizioni, tempo e grado di certezza, prima di accettare il tono. Possiamo chiedere una versione più leggibile: «Se i dati arrivano entro mercoledì, potrei consegnare venerdì». La condizione resta visibile.

La variante è un messaggio in cui la persona vuole esprimere disaccordo senza irrigidire il rapporto. Se il modello elimina il disaccordo per rendere il testo gentile, non ha svolto il compito. Riusa l'assistenza per trovare una forma, poi rileggi come destinatario e come autore: ciò che stai per dire è ancora tuo, anche negli impegni e nei limiti?

L'IA che non sembra IA

Immaginiamo robot, assistenti virtuali e grandi innovazioni, ma la maggior parte degli incontri è ordinaria e silenziosa. Completamento automatico, trascrizione vocale, filtro antispam, una playlist che conosce i tuoi gusti, una notifica che invita ad alzarti: possono usare modelli di IA o altre forme di automazione, secondo il servizio. Non sono spettacolari, ma plasmano il comportamento.

Più l'IA si integra senza attrito, più deleghiamo anche la fatica. Non ci fermiamo a cercare, confrontare, riscrivere o considerare: il sistema lo fa per noi. È comodo, ma non neutrale.

La pausa che scompare

La pratica Zen onora la pausa: il respiro prima del tiro con l'arco, il silenzio prima della frase, lo spazio fra pensiero e azione. Progettata per l'efficienza, l'IA può consumare questa pausa. Riempie lo spazio e completa una frase prima che l'abbiamo sentita. Rischiamo di agire automaticamente.

Esiste una differenza sottile fra ricevere assistenza ed essere trasportati, fra usare uno strumento e diventarne un'estensione. Se l'app delle mappe cambia percorso prima che valutiamo il traffico, perdiamo un atto di giudizio. Se riscrive un messaggio, potremmo non notare il nuovo tono. Se un algoritmo nasconde contenuti e ne privilegia altri, potremmo ignorare ciò che abbiamo perso.

Il filtro può diventare parte della nostra visione del mondo senza presentarsi come una decisione. Se vediamo sempre le fotografie selezionate dal programma, gli articoli ordinati dal servizio o le formulazioni suggerite dalla tastiera, possiamo perdere il confronto con ciò che è stato escluso. La comodità rimane un vantaggio, ma il suo valore dipende anche dalla possibilità di riconoscere e cambiare quella selezione.

L’intelligenza sullo sfondo

Questi strumenti incorporano scelte progettuali, supposizioni di addestramento, ottimizzazioni e valori. Ogni piccola delega, dal controllo grammaticale al riconoscimento facciale, porta una catena di scelte invisibili. Chiedendo di ordinare, suggerire, rispondere e stabilire priorità, invitiamo un modello nella coreografia intima della vita quotidiana.

Come nello Zen anche spazzare il pavimento è una pratica, così l’uso dell’IA deve diventare consapevole. Non drammatico: presente.

Vivere con la macchina

Essere umani con l’IA significa a volte semplicemente accorgersene. Osservare come ci orienti, quando interrompa e come cambi ciò a cui prestiamo attenzione. Questo strumento mi aiuta ad agire con più cura? A vedere meglio? Oppure rende lisce cose che dovrei sentire?

Un monaco Zen porta l’acqua consapevole del suo peso. Se l’IA porta qualcosa che tenevamo noi, anche la consapevolezza deve seguirlo. La sua presenza non va ignorata o drammatizzata. L’IA quotidiana è reale, vicina e può essere plasmata attraverso attenzione, uso e scelta.

Momenti concreti che potremmo non aver notato

Immaginiamo spesso l’IA lontana, fra assistenti robotici e laboratori futuristici. È già qui, silenziosamente.

Scrittura automatica dei messaggi. Cominci con «Ciao, ti scrivo per…» e un programma di posta suggerisce «…sapere come stanno andando le cose». Un sistema di completamento può proporre una continuazione usando schemi appresi. Fa risparmiare tempo e conserva coerenza di tono; può però far adottare inconsapevolmente formule impersonali.

Organizzazione delle fotografie. Cerchi «spiaggia 2021» e trovi subito le immagini, senza etichette o cartelle manuali. L’IA riconosce schemi in conte-

nuto, data e posizione. Organizza i ricordi senza fatica, ma possiamo dipendere dalla selezione della macchina: e se escludesse qualcosa di significativo?

Suggerimenti di prodotti che ci conoscono troppo bene. Guardi uno zaino e ti segue per giorni su siti diversi. Un sistema analizza il comportamento, lo confronta con utenti simili e stima ciò che potrebbe spingerti all'acquisto. Aiuta a scoprire oggetti pertinenti, ma una personalizzazione eccessiva può limitare le scelte e favorire acquisti impulsivi.

Calendari e promemoria predittivi. «È ora di partire per l'appuntamento delle 15. Il traffico è scorrevole». Il calendario si è collegato a mappe e abitudini, prevedendo spostamenti e orario di partenza. Può ridurre stress e ritardi; la dipendenza può indebolire memoria e stima personale dei tempi.

Riformulare i pensieri. Un programma propone «Vuoi riscrivere per maggiore chiarezza?». Suggerisce una versione concisa che può cambiare intenzione o sfumature. Accetti e prosegui. Migliora leggibilità e tono, ma rischi di perdere la tua voce in favore della levigatura algoritmica.

Questi sistemi sussurrano. Nel sussurro plasmano pensiero, azione e relazioni.

> «Quando lo strumento diventa invisibile, il gesto deve restare consapevole.» Riflessione proposta nel libro.

La cerimonia del tè aumentata

Nella scena che accompagna il capitolo una ciotola, una frusta e un bollitore sono disposti con cura, entro uno spazio percorso da tracce circolari. Il tablet e il cavo del bollitore appartengono allo stesso insieme visivo: il digitale non entra come un elemento spettacolare, ma come qualcosa che abbiamo imparato a lasciare vicino. L'illustrazione rende leggibile il rapporto tra gli oggetti senza pretendere di descrivere una cerimonia tradizionale.

È questo il punto. Il cambiamento più potente avviene quando l'IA entra nella vita quotidiana: l'assistente invisibile in tasca, nel programma di scrittura e in casa. L'immagine normalizza la tecnologia, presente e discreta, integrata nel rituale senza sostituirlo interamente.

Questa fusione ha significato. Nello Zen conta come spazzi e versi il tè, perché conta la tua presenza. Immagina ora un sistema che versa il tè, ottimizza la temperatura dell'acqua e scrive la risposta prima che tu abbia sentito ciò che volevi dire. È comodità e anche una sottile erosione della pausa.

L'immagine invita a chiedersi se siamo pienamente presenti o trasportati, se lo strumento ci renda più presenti o soltanto più veloci, se il gesto sia ancora nostro.



Figura 13 · La cerimonia del tè e gli oggetti digitali condividono lo stesso spazio

La figura conserva l'incontro fra ciotola, frusta, tazze, tablet e bollitore con cavo. Non pretende di descrivere una cerimonia tradizionale corretta. È una scena proposta nel libro: il digitale entra nel gesto ordinario, e ci chiede dove si trovi ancora la nostra attenzione.

> **Automazione non significa sempre IA.** Una notifica può usare una regola fissa, un calendario può eseguire un calcolo e un termostato può seguire una soglia. Per parlare di modello appreso occorre sapere come sia costruito il servizio. Dal solo comportamento nell'interfaccia non possiamo sempre dedurlo.

La pausa prende così una forma concreta: rileggere una condizione prima di accettare un suggerimento, cercare una fotografia esclusa dal filtro, controllare il motivo di un promemoria. Lo strumento può togliere fatica mentre noi conserviamo il senso del gesto. Nel capitolo successivo questa attenzione ci aiuterà anche a scegliere quanto calcolo serve e dove convenga svolgerlo.

Capitolo 9 Piccolo è bello intelligenza vicina alle persone

Quando si parla di IA, le dimensioni diventano facilmente una misura del valore: più parametri, più dati, più calcolo sembrano indicare una direzione inevitabile. La scala ha reso possibili capacità importanti, ma non risponde da sola alla domanda che ci interessa nella vita quotidiana: quale sistema è sufficiente per questo bisogno, in questo luogo e con questi vincoli? Il capitolo precedente ci ha portato negli oggetti ordinari; ora possiamo guardarli da vicino senza presumere che ogni funzione richieda la macchina più grande disponibile.

Immaginiamo un filo d'erba come occasione per osservare la Via, il percorso di pratica e comprensione, e un monaco che porta soltanto gli strumenti necessari. Leggiamo queste immagini come un invito alla sufficienza. Non celebrano la scarsità né dimostrano che un modello piccolo sia sempre migliore: suggeriscono di riconoscere ciò che il compito richiede e ciò che aggiungiamo per abitudine. La semplicità ha valore quando rende possibile un lavoro preciso e lascia più chiaro il rapporto con le persone.

Torniamo al termostato. «Accendi sotto 18 °C» è una regola fissa: confronta la temperatura con la soglia, senza aver bisogno di un modello appreso. Se vogliamo invece prevedere quando la stanza raggiungerà la temperatura desiderata, combinando temperatura esterna, abitudini e andamento precedente, un modello può diventare utile. Dobbiamo confrontare la sua previsione con ciò che accade davvero e chiedere se migliori una soluzione più semplice. Un dispositivo indossabile che misura passi o frequenza cardiaca e una telecamera che osserva un incrocio offrono altri compiti circoscritti; la loro vicinanza alle persone non garantisce accuratezza, proporzione o rispetto dei dati.

Il calcolo nel luogo in cui serve

Con *edge AI* indichiamo un'elaborazione vicina alla sorgente dei dati, spesso sul dispositivo che li raccoglie. Può avvenire in una casa, in un oggetto indossabile, in un sensore lungo una strada o in un piccolo computer. Invece di inviare ogni informazione a un centro remoto per ottenere una risposta, il sistema svolge localmente una parte del lavoro. La prossimità riguarda il calcolo e il contesto, non una qualità spirituale della macchina.

Sul termostato possiamo distinguere tre domande. Dove viene calcolata la previsione? Se avviene nel dispositivo, l'elaborazione è locale. Quanto è grande il modello? Questa è una scelta separata: un modello piccolo può anche funzionare su un server lontano. Quali dati escono dalla stanza? Persino un calcolo locale può essere accompagnato dall'invio di registri all'esterno. Separare luogo, dimensione e trasferimenti ci permette di valutare il progetto senza attribuire alla parola «edge» tutte le qualità desiderate.

Questa scelta può ridurre la dipendenza dalla connessione, il trasferimento di dati e alcuni tempi di attesa. Può anche permettere funzioni quando la rete non è disponibile, se tutte le risorse necessarie sono locali. Tuttavia il tempo di calcolo rimane, il dispositivo ha limiti di memoria ed energia e altre funzioni dell'app possono ancora sincronizzare informazioni. I vantaggi devono essere verificati nel percorso completo, non dedotti dalla sola parola «locale».

> **Vicino non significa riservato per definizione.** Un modello può elaborare sul dispositivo mentre l'app invia registri, aggiorna archivi o sincronizza risultati. Occorre sapere quali informazioni vengono trattate, dove vengono conservate e quali trasferimenti sono attivi. Allo stesso modo, un modello piccolo non è automaticamente interpretabile: dimensioni e leggibilità del comportamento sono questioni distinte.

L'autonomia richiede condizioni concrete

L'idea di eseguire un modello vicino a noi può aumentare il controllo sul servizio, soprattutto quando disponiamo dei componenti, delle licenze e della documentazione necessari. Possiamo ridurre alcuni vincoli di un'infrastrut-

tura remota, decidere quando aggiornare e conservare una funzione anche durante un'interruzione di rete. Ma possedere un dispositivo non significa comprendere tutto ciò che esegue, e una licenza aperta non sostituisce la capacità di mantenere il sistema.

Un computer su una sola scheda può essere il punto di partenza per una stazione meteorologica, un traduttore, un assistente o un tutor. Sono possibilità da valutare rispetto al compito e all'hardware. Un tutor che lavora su materiali limitati può avere esigenze diverse da un servizio che spiega documenti lunghi in molte lingue; un organizzatore di documenti sanitari non va confuso con una diagnosi affidabile. La vicinanza offre una scelta tecnica, non un'approvazione universale degli usi.

Anche la sostenibilità richiede un confronto reale. Modelli più piccoli possono richiedere meno risorse per alcune esecuzioni, ma il consumo dipende dalla frequenza d'uso, dall'hardware e dal lavoro complessivo. Se un sistema viene attivato inutilmente molte volte, il vantaggio del singolo calcolo può ridursi. La misura utile mette a confronto soluzioni che svolgono il medesimo compito, includendo trasferimenti e manutenzione quando contano.

Sedere ai piedi della montagna



Figura 14 · La prossimità come rapporto con il contesto

La figura mostra una persona ai piedi di una montagna, piccola davanti a un paesaggio ampio. Non sta scalando, misurando o conquistando: osserva da una posizione che le permette di incontrare il luogo. La montagna evoca la complessità e la scala di ciò che cerchiamo di comprendere; la figura seduta evoca una relazione in cui la conoscenza può cominciare dalla vicinanza, senza coincidere con il dominio.

Questa immagine non spiega un'architettura edge, ma chiarisce l'orientamento della domanda. Per il sensore, stare nel luogo significa ricevere dati del contesto; per chi progetta, significa conoscere le persone e le condizioni d'uso. Le due prossimità non sono intercambiabili: un dispositivo può essere vicino e

il progetto rimanere lontano dai bisogni. Il valore nasce quando il luogo del calcolo e la comprensione del compito si incontrano.

Scegliere la misura del bisogno

Torniamo alla stanza in cui vogliamo rilevare una temperatura troppo alta. Prima di scegliere un modello, proviamo la soluzione semplice: sensore, soglia e segnalazione. Se la richiesta diventa prevedere l'andamento usando occupazione e condizioni esterne, definiamo quale anticipo sia utile e quale errore sia accettabile, poi confrontiamo le alternative. La prova osserva qualità della previsione, tempo di risposta, funzionamento senza rete e consumo sullo stesso compito.

La variante cambia il lavoro: il dispositivo deve spiegare documenti lunghi in lingue diverse. Il sistema locale può non avere risorse o qualità sufficienti; quello remoto può richiedere trasferimenti incompatibili con i dati. Possiamo ridurre la funzione, dividere il lavoro o rinunciare a una parte dell'automazione. La scelta nasce dai vincoli osservati, non dalla fedeltà a un'etichetta.

Il criterio trasferibile è la sufficienza: scegliere ciò che soddisfa il bisogno e che sappiamo mantenere e verificare. «Piccolo è bello» rimane così una domanda aperta sul rapporto tra capacità e misura, invece di diventare un altro assoluto. Il percorso della parte successiva comincia da qui, nel momento in cui la nostra attenzione decide quale lavoro chiedere alla macchina e quale possibilità lasciare alle persone.

Parte IV La pratica dell’attenzione



La montagna, l'impronta e l'onda evocano un luogo in cui fermarsi, il segno di un passo e il movimento che continua intorno a noi. La quiete della montagna non annulla il cammino. Offre una posizione da cui riconoscere dove siamo e quale direzione stiamo seguendo, invece di confondere il movimento con la scelta.

Dopo aver osservato capacità, cultura e abitudini, possiamo tornare all'interazione con una domanda più concreta: quale lavoro vogliamo chiedere al sistema e come conserviamo il giudizio mentre lo svolge? La presenza prende forma in una richiesta delimitata, nella lettura di una risposta e nella possibilità di cambiare o interrompere una delega.

Seguiremo la preparazione di un incontro pubblico: chiariremo che cosa i partecipanti dovrebbero imparare, confronteremo due programmi e controlleremo una pagina informativa. Il caso offre un terreno su cui provare un criterio e poi trasferirlo a una lettera o a una lezione. Le domande aperte rimangono con noi, ma non ci impediscono di agire: ci aiutano a sapere perché scegliamo una proposta e quale responsabilità conserviamo.

Capitolo 10 Fermarsi prima di chiedere

Il cursore lampeggia e la domanda sembra pronta: «Che cosa dovrei fare?». È una frase breve, ma può contenere un lavoro intero lasciato fuori dallo schermo. Quale situazione stiamo affrontando? Che cosa possiamo cambiare? A che cosa non vogliamo rinunciare? Un modello può riempire quello spazio con una risposta ordinata. Il fatto che lo spazio sia stato riempito non significa che sia stato compreso.

Nei capitoli precedenti abbiamo distinto il contesto dalla conoscenza, il tono dai fatti e la comodità dalla delega. Ora possiamo usare queste distinzioni prima della risposta. La pausa non è un rituale da aggiungere a ogni clic. È il momento in cui decidiamo quale lavoro chiedere allo strumento e quale lavoro rimane nostro.

Una domanda troppo grande

Immaginiamo una piccola associazione che voglia organizzare un incontro pubblico. Chi la coordina apre un chatbot e scrive: «Organizza un evento di successo». Il modello può proporre una sede, una campagna sui social, ospiti autorevoli e un programma pieno. È una risposta plausibile e quasi impossibile da valutare. Non sappiamo che cosa significhi successo: molte presenze, un confronto utile, nuove iscrizioni o un servizio per chi non partecipa mai?

La soluzione abituale è aggiungere istruzioni. Ma una richiesta lunga può continuare a nascondere lo stesso problema. Se non abbiamo deciso lo scopo, i dettagli amplificano una direzione ancora incerta. Prima di cercare un prompt migliore, serve una descrizione migliore della situazione.

Proviamo a scrivere, fuori dalla chat: «Vorremmo aiutare venti persone del quartiere a capire come proteggere i propri account. Abbiamo una sala per due ore e un volontario che sa spiegare le basi. Non raccoglieremo password o

dati personali». Sono poche frasi. Rendono visibili destinatari, risorse e confini.

Da qui nasce una richiesta più utile: «Proponi due programmi per questo incontro. Per ogni programma spiega come un partecipante potrà provare di aver imparato a riconoscere un messaggio sospetto. Non proporre strumenti a pagamento. Se mancano informazioni, chiedile prima di scegliere». Non chiediamo una celebrazione del progetto. Chiediamo alternative che possano essere confrontate.

Chiarire che cosa stiamo delegando

Il modello può distribuire tempi, formulare esempi e confrontare due sequenze. Non può conoscere la sala se non gliela descriviamo, verificare la disponibilità dei volontari senza informazioni, o stabilire da solo che cosa sia rispettoso per quel gruppo. Il compito deve restare commisurato al contesto disponibile.

> **Scopo, compito e criterio.** Lo scopo è il cambiamento desiderato: i partecipanti riconoscono un messaggio sospetto. Il compito è il lavoro affidato al modello: proporre due programmi. Il criterio è ciò con cui giudichiamo il risultato: ogni programma comprende una prova accessibile entro due ore. Confondere questi tre livelli rende facile accettare un testo ben scritto che non serve allo scopo.

Una buona domanda può anche chiedere quali informazioni manchino. Questo non garantisce che il modello individui tutte le lacune. Aiuta però a farle emergere, purché le leggiamo come proposte da controllare. Se il sistema domanda l'età media dei partecipanti ma ignora l'accessibilità della sala, quella mancanza resta nostra da vedere.

Prima di inviare il contesto

Per ottenere precisione siamo tentati di condividere tutto. L'elenco dei partecipanti, i messaggi privati, le difficoltà personali. Ma il bisogno del modello non coincide con tutto ciò che sappiamo. Per costruire il programma bastano, per

esempio, il livello iniziale, le esigenze di accessibilità e il numero previsto. Non servono nomi e storie individuali.

Il criterio è concreto: chiediamo se ogni dettaglio cambia davvero la proposta. Quando non la cambia, possiamo ometterlo. Quando la cambia, possiamo spesso descrivere la condizione senza identificare una persona. «Due partecipanti useranno uno schermo con caratteri grandi» offre un vincolo di progetto; una cartella con i loro dati sanitari offrirebbe molto di più del necessario.

La scelta dipende anche dal servizio: dove vengono trattati i dati, chi può accedervi, che cosa viene conservato? Questo libro non presume una risposta uguale per ogni prodotto. Prima di condividere materiali riservati, occorre conoscere le condizioni effettive del sistema usato.

Una risposta da confrontare

Supponiamo che la prima proposta dedichi novanta minuti alla spiegazione e dieci alla prova. La seconda alterni brevi spiegazioni e tre messaggi di esempio. Il confronto non si decide in base al tono. Proviamo il secondo percorso con il volontario: riesce a spiegare il primo caso in dieci minuti? Un partecipante può motivare che cosa lo rende sospetto? Il gruppo ha tempo per fare domande?

Nel nostro incontro, la prova può essere molto semplice: dopo una breve spiegazione mostriamo due messaggi fittizi, uno ordinario e uno sospetto, e chiediamo di motivare la scelta. Se una persona risponde correttamente soltanto perché riconosce l'esempio già visto, cambiamo mittente e formulazione. Così la preferenza per un programma nasce da un comportamento osservabile, senza dover credere alla sua presentazione più convincente.

Se la prova mostra che il programma è troppo fitto, restituiamo al modello un dato preciso: «Il primo esercizio ha richiesto venti minuti. Riduci il numero dei casi e conserva la discussione degli errori». Abbiamo trasformato una pre-

ferenza vaga in un'osservazione utile. La revisione nasce da ciò che è accaduto, non dal desiderio di avere una risposta più elegante.

Il percorso segue il movimento: riconoscere lo scopo, delimitare la delega, chiedere una proposta, provarla e decidere. Il ritorno alla domanda ricorda che una prova può cambiare la domanda iniziale.

Quando conviene non chiedere

A volte il lavoro necessario è una conversazione con una persona. Se l'associazione è divisa sullo scopo dell'incontro, produrre cinque programmi può rinviare quella conversazione. Il modello aiuta a formulare alternative; non conferisce legittimità a una scelta che gli interessati non hanno discusso.

Altre volte abbiamo già una procedura semplice e affidabile. Se la sala va prenotata compilando un modulo di cinque campi, un sistema generativo può aggiungere passaggi senza migliorare il risultato. Fermarsi significa anche riconoscere quando basta lo strumento che abbiamo.

Portare la pausa altrove

Sostituiamo l'incontro pubblico con la preparazione di una lettera a un collega. Lo scopo potrebbe essere chiarire una responsabilità; il compito, proporre una formulazione rispettosa; il criterio, conservare fatti e impegni senza attribuire intenzioni non espresse. Una richiesta come «Scrivi una risposta vincente» non soddisfa questo bisogno. Una richiesta che distingua osservazioni, richiesta concreta e spazio per il confronto può farlo.

Prima di inviarla, prova a dire quale frase dovrebbe restare vera anche se il modello cambiasse completamente lo stile. Quella frase individua il terreno della domanda. Se non riesci a formularla, può essere utile fare ancora un passo fuori dalla chat.

La pausa non ci allontana dal lavoro. Ci permette di entrare nel lavoro sapendo che cosa chiediamo.

Capitolo 11 Leggere una risposta senza aggrapparsi

Una risposta arriva. Ha un titolo, tre argomenti e una conclusione. La forma è rassicurante: qualcuno sembra avere già separato ciò che conta da ciò che possiamo lasciare da parte. Dopo il lavoro di formulare la domanda, è naturale desiderare questa chiarezza. Ma la risposta diventa utile quando possiamo attraversarla, non soltanto accoglierla.

Il capitolo sulle allucinazioni ha mostrato perché il tono non certifichi i fatti. Qui affrontiamo un problema ulteriore: anche una risposta corretta può chiudere troppo presto una scelta. Può selezionare criteri che non abbiamo condiviso, ignorare una persona coinvolta o trasformare una possibilità in una conclusione inevitabile.

Tre livelli nella stessa pagina

Immaginiamo di aver chiesto quale dei due programmi dell'incontro sia preferibile. Il modello risponde: «Il secondo è migliore perché gli adulti apprendono sempre più velocemente facendo esercizi. Richiede tre gruppi da dieci persone e porterà a una partecipazione del novanta per cento».

Leggiamo separatamente le tre affermazioni. «Tre gruppi da dieci» è una proposta organizzativa: $3 \times 10 = 30$, mentre aspettiamo venti persone. «Gli adulti apprendono sempre più velocemente facendo esercizi» è un'affermazione generale sull'apprendimento; richiede evidenze e condizioni che la frase non indica. «Partecipazione del novanta per cento» è una previsione: senza dati su inviti, adesioni o prove precedenti, non sappiamo da dove venga il numero. Il controllo cambia per ciascuna frase, anche se il modello le ha unite in un solo consiglio.

La prima operazione è separare. Possiamo conservare l'idea di alternare spiegazione e pratica, correggere il numero dei gruppi e rifiutare la percentuale

non sostenuta. L'alternativa non è credere a tutto oppure buttare tutto. È attribuire a ciascuna frase il lavoro che può legittimamente svolgere.

> **Fatto, inferenza, proposta.** Un fatto viene dichiarato sul mondo e richiede un riscontro. Un'inferenza collega informazioni a una conclusione e va controllata nei passaggi. Una proposta offre un'azione e va giudicata rispetto a scopo e vincoli. La stessa risposta può contenere tutti e tre. Una proposta interessante non rende vero il fatto che la accompagna.

Nel caso dell'incontro, «Sono previste venti persone» è un fatto da confrontare con le adesioni. «Serviranno due gruppi se ogni facilitatore può seguirne dieci» è un'inferenza: usa quel numero e una capacità che dobbiamo confermare. «Alterniamo spiegazione ed esercizio» è una proposta da provare rispetto allo scopo. Dare un nome ai tre lavori rende più facile scegliere il controllo, senza assegnare una fiducia unica alla risposta intera.

Rallentare dove il testo accelera

Parole come «sempre», «inevitabile», «ottimale» o «dimostrato» segnalano un punto da guardare. Non provano un errore; chiedono quale evidenza sostenga una formulazione così forte. Un titolo preciso può nascondere una fonte inesistente, come abbiamo visto. Una fonte esistente può non sostenere la conclusione per cui viene citata.

Quando una risposta indica un articolo, apriamo il riferimento. Controlliamo titolo, autori, data e soprattutto il passaggio pertinente. Se il testo è disponibile soltanto in abstract, la verifica resta limitata a ciò che l'abstract dice. Non estendiamo una frase breve a risultati o condizioni che non abbiamo letto.

Chiedere al modello «Sei sicuro?» può produrre una correzione o una conferma convincente. Entrambe sono nuove uscite dello stesso sistema. Per uscire dal circuito serve un riscontro diverso: il documento, il dato di partenza, un calcolo o una prova osservabile. Una seconda risposta non è automaticamente una seconda fonte indipendente.

Una verifica che cambia la lettura

Torniamo ai treni del capitolo 4. La soluzione dice che il secondo raggiunge il primo dopo due ore dalla propria partenza. Per verificarla non occorre interrogare un altro modello. Nel momento proposto il secondo ha percorso 90 × 2 = 180 chilometri. Il primo viaggia da tre ore e ha percorso 60 × 3 = 180 chilometri. Le distanze coincidono. Abbiamo controllato una condizione necessaria del risultato.

La verifica non finisce al numero. I treni devono partire dallo stesso punto, muoversi nella stessa direzione e mantenere velocità costanti. Se un'informazione manca, la soluzione è valida sotto queste ipotesi, non universalmente. È il modo in cui una risposta apparentemente semplice torna ad avere confini.

La verifica non assegna una luce verde al testo intero. Ogni affermazione segue il proprio percorso: un fatto verso una fonte, un calcolo verso le operazioni, una proposta verso una prova e le persone coinvolte.

Il piacere di una risposta può restare

Non aggrapparsi non significa diventare sospettosi di ogni frase. Un'immagine poetica può aprire una prospettiva anche quando non descrive un fatto. «Il silenzio non risponde: accoglie», nel racconto illustrativo del capitolo 4, funziona come immagine letteraria. Non serve chiedere quale esperimento dimostri che il muschio accolga una bambina. Serve riconoscere che stiamo leggendo una storia.

La stessa frase diventerebbe inadeguata se usata come risposta a una persona che chiede un aiuto concreto urgente. Il contesto decide quale lavoro devono fare le parole. Una buona lettura tiene insieme apprezzamento e discernimento: vede ciò che il testo offre e ciò che non può offrire.

Anche il disaccordo merita una pausa. Se una risposta contraddice ciò che credevamo, possiamo essere tentati di cercare una nuova formulazione che ci dia ragione. Prima di cambiare il prompt, chiediamo che cosa potrebbe confutare

la nostra posizione. La verifica deve poter cambiare anche noi, non soltanto il testo.

Conservare le alternative

Per una decisione, una buona risposta dovrebbe rendere visibili opzioni e costi. Se il programma con più esercizi richiede due facilitatori e ne abbiamo uno, può essere impraticabile. Se quello con una lezione lunga è più semplice da organizzare ma lascia poco spazio alle domande, possiamo ridurlo e aggiungere una sola prova. La scelta migliore emerge dall'adattamento, non dalla fedeltà a una proposta iniziale.

Possiamo chiedere al modello una critica della propria proposta, ma dobbiamo fornire il criterio: «Quale parte fallisce se arriva un gruppo che non usa abitualmente un computer?». È più utile di una generica richiesta di autocritica. La risposta resta da verificare, ma incontra un limite riconoscibile.

Una variante da attraversare

Supponi ora di leggere una risposta che confronta due strumenti per organizzare fotografie. Contiene una funzione descritta, una promessa di riservatezza e un consiglio sul prezzo. Quale riscontro serve per ciascuno? La funzione richiede una prova o documentazione; la riservatezza dipende dalle condizioni reali e dalle impostazioni; il prezzo va controllato alla data dell'acquisto. Una sola recensione non sostituisce questi tre controlli.

Se una parte non è verificabile, puoi rimandare la decisione, ridurre la delega o scegliere una soluzione conosciuta. La mancanza di una risposta finale è a volte il risultato più corretto della lettura.

Il testo è arrivato. Il lavoro del giudizio comincia qui.

Capitolo 12 Collaborare conservando il proprio giudizio

Una persona e un modello possono produrre insieme un buon documento. Per capire se la collaborazione funziona, tuttavia, non basta guardare il documento. Occorre guardare anche la persona: sa spiegare le scelte? Riconosce un errore? Riesce a continuare quando lo strumento non è disponibile?

Abbiamo imparato a delimitare una richiesta e a leggere criticamente una risposta. Il passo successivo è distribuire il lavoro senza perdere la capacità di comprenderlo. L'intelligenza aumentata ha senso quando aumenta ciò che possiamo fare e valutare. Se il risultato cresce mentre il giudizio si restringe, la collaborazione merita di essere ripensata.

Una bozza che appartenga a qualcuno

Immaginiamo di dover scrivere una pagina informativa per l'incontro del quartiere. Il modello può organizzarla rapidamente: titolo, destinatari, programma, iscrizione. Il coordinatore potrebbe pubblicarla senza rileggerla, oppure correggere soltanto i refusi. Entrambe le scelte lasciano aperta una domanda: chi sa rispondere del contenuto?

Proviamo un altro percorso. Prima della chat, il coordinatore scrive i fatti essenziali con parole proprie: data, luogo, durata, obiettivo, accessibilità e modalità d'iscrizione. Chiede poi al modello una versione chiara per lettori non specialisti, specificando di non aggiungere informazioni. Confronta la bozza con quei fatti e controlla ogni promessa. «Imparerai a proteggerti da ogni truffa» va corretto: un incontro breve può insegnare segnali e procedure utili, non garantire protezione assoluta.

Se la bozza cambia «Ingresso gratuito, iscrizione obbligatoria» in «Ingresso libero», il coordinatore può rifiutare quella modifica: il lettore potrebbe credere che non serva iscriversi. Può invece accettare «La partecipazione è

gratuita; per prenotare il posto occorre iscriversi», che rende la condizione più esplicita. Il controllo riguarda ciò che una persona farà dopo aver letto, oltre alla somiglianza delle parole.

Il modello ha svolto un lavoro reale sulla forma. La persona ha conservato la responsabilità del significato. La distinzione resta visibile perché esiste un punto di partenza con cui confrontare la bozza.

> **Autonomia e controllo effettivo.** Essere presenti davanti allo schermo non significa avere controllo. Per intervenire servono informazioni comprensibili, tempo, competenza e la possibilità concreta di cambiare o interrompere il risultato. Un pulsante «conferma» non basta se tutto spinge ad accettare senza leggere. [R10]

Rendere visibile la collaborazione

Quando lavoriamo in gruppo, la provenienza delle decisioni conta. Non occorre trasformare ogni frase in un verbale. Può bastare conservare le fonti dei fatti, una nota sulle parti generate e il nome di chi ha rivisto il testo. Se nasce una contestazione, sappiamo quale passaggio tornare a controllare.

La trasparenza deve servire al destinatario. Per una pagina pubblica, può essere utile indicare che il testo è stato preparato con assistenza automatica e rivisto dall'associazione, soprattutto quando le persone potrebbero attribuirgli un'autorevolezza diversa. Per un testo interno di lavoro, conta anche rendere visibili le parti ancora da verificare. Le convenzioni vanno definite prima, affinché un collaboratore non scambi una bozza per una decisione.

Il contributo umano non si misura dal numero di parole riscritte. Una sola correzione può impedire una promessa falsa; cento modifiche stilistiche possono lasciare intatto l'errore. La domanda è quali scelte siano state comprese e controllate.

Non chiedere una relazione che il sistema non può sostenere

Un chatbot può parlare con gentilezza e continuità. Questo può essere piacevole e aiutare a formulare pensieri. La sensazione di essere ascoltati è un'esperienza dell'utente; non dimostra una reciproca partecipazione emotiva della macchina.

Il problema compare quando la disponibilità del sistema sostituisce ogni confronto esterno. Se il modello conosce soltanto il nostro racconto, può confermare una versione incompleta di un conflitto. Se gli chiediamo di spiegare il comportamento di un collega, può attribuire motivazioni che non ha modo di conoscere. Una collaborazione prudente usa la chat per preparare domande o chiarire osservazioni, poi torna alle persone coinvolte.

Per esempio, invece di chiedere «Perché il collega mi manca di rispetto?», possiamo scrivere: «Aiutami a distinguere fatti osservati, mie interpretazioni e domande da porre in un confronto». Il risultato non determina chi abbia ragione. Rende più facile una conversazione che rimane umana.

Il controllo cresce attraverso piccole prove

Un'abitudine utile consiste nel fermarsi dopo una parte della risposta e spiegarla senza guardare lo schermo. Se non sappiamo dire perché una scelta sia stata fatta, non siamo ancora pronti a delegarne gli effetti. Possiamo chiedere una spiegazione più chiara, ma dobbiamo poi verificarla con un esempio diverso.

Nel programma dell'incontro, abbiamo scelto di alternare spiegazione ed esercizio per rendere osservabile l'apprendimento. Sapremmo adattare questa scelta a un incontro di trenta minuti? Potremmo conservare un caso, una breve discussione e una prova finale. Se ci limitiamo a ridurre proporzionalmente tutti i tempi, rischiamo di rendere inutilizzabile ogni attività. Comprendere il motivo della struttura permette di modificarla.

Non serve fare questa prova per ogni messaggio ordinario. Conviene quando il risultato contiene una decisione importante, un procedimento nuovo o una competenza che vogliamo sviluppare. La quantità del controllo deve seguire la conseguenza dell'errore.

Dalla collaborazione individuale a quella organizzativa

Un'organizzazione ha bisogno di sapere chi può usare quale sistema, con quali dati e per quali compiti. Queste scelte non diventano superflue perché il prodotto è facile da usare. Il coordinatore che ha controllato la pagina informativa non deve automaticamente poter caricare l'archivio dei partecipanti o inviare messaggi a nome di tutti.

Definiamo quindi una delega proporzionata. Il sistema può preparare una bozza; l'invio richiede una persona responsabile. Può confrontare versioni; una modifica ai fatti richiede conferma. Può segnalare una lacuna; non deve riempirla silenziosamente con un'invenzione. Sono confini pratici che trasformano un'intenzione etica in un comportamento del servizio.

Quando qualcosa va male, la domanda non è soltanto chi abbia premuto un pulsante. Occorre capire se l'interfaccia rendeva il controllo possibile, se il gruppo aveva una procedura e se qualcuno poteva contestare il risultato. La responsabilità segue tutta la catena.

Una prova di indipendenza

Immagina che domani il modello non sia disponibile. Sapresti aggiornare la pagina dell'incontro, mantenendo i fatti e spiegando le scelte? Se sì, la collaborazione ha lasciato anche una capacità. Se no, individua il passaggio che rimane opaco e ricostruiscilo con gli altri.

In un contesto diverso, come la preparazione di una lezione o la revisione di un progetto, il criterio rimane lo stesso: delegare il lavoro che lo strumento svolge bene, conservare la capacità di verificarlo e sapere quando chiamare

un'altra persona. Un futuro in cui la tecnologia sostiene le relazioni prende forma in questi gesti. Una tecnologia che aiuta non deve rendere meno visibili le persone a cui dovrebbe servire.

Parte V Progettare con presenza



Le mani aperte e il loto ci invitano a riflettere sul gesto della progettazione. Le mani offrono e, nello stesso tempo, lasciano spazio; il loto mantiene una relazione con il fondo da cui cresce, pur aprendosi sulla superficie dell'acqua. Possiamo leggere queste immagini come un invito a costruire senza dimenticare il contesto e senza occupare tutto lo spazio delle persone a cui il sistema dovrebbe servire.

L'IA centrata sull'essere umano porta questa attenzione nelle scelte di progetto. Non basta che una funzione sia rapida o piacevole: occorre capire chi possa usarla, chi ne subisca gli errori e chi abbia modo di correggerla. La cura deve diventare un comportamento osservabile dell'interfaccia, una regola della delega e una risorsa dell'organizzazione.

In questa parte una risposta con l'orario sbagliato ci porterà dai principi alle scelte di progetto. Vedremo chi può segnalare l'errore, chi lo corregge e come provare che il servizio aiuti davvero persone diverse. Un laboratorio mostrerà poi come insegnare a riconoscere una modifica utile e una che cambia il significato. Il futuro evocato dal libro comincia in queste possibilità concrete di partecipazione e intervento.

Capitolo 13 L'IA centrata sull'essere umano

> «Per costruire per gli altri, impara prima a sederti con loro.» Riflessione proposta nel libro.

Immaginiamo che l'assistente dell'incontro risponda «Cominciamo alle 17», mentre l'avviso aggiornato dice «Ore 18». La persona vede la discrepanza: deve poter aprire la fonte, capire chi la aggiorna e segnalare l'errore a qualcuno che possa correggerlo. Se il servizio offre soltanto una risposta elegante, lascia alla persona il problema senza darle gli strumenti per risolverlo. Da questo caso nasce una domanda di progetto: per chi è utile l'automazione e quale possibilità d'azione conserva chi la incontra?

L'IA centrata sull'essere umano, detta HCAI dalla sigla inglese *human-centred artificial intelligence*, parte da domande come questa. La persona partecipa alle scelte, porta il proprio contesto, la propria storia e dignità e può contribuire al progetto. Il campo intreccia informatica, progettazione, psicologia, etica e discipline umanistiche: capire il calcolo non basta se non comprendiamo anche la situazione di chi userà il servizio. Si cercano sistemi comprensibili e in sintonia con i valori delle persone, oltre che facili da usare.

Come lo Zen, l'HCAI comincia dalla presenza: progettazione radicata nella consapevolezza, nell'ascolto e nella comprensione del rapporto umano con una macchina, anziché automazione fine a se stessa.

Che cosa è l'IA centrata sull'essere umano

È un approccio alla progettazione, allo sviluppo e all'impiego dell'IA che privilegia bisogni, valori e capacità d'azione delle persone in ogni fase del ciclo di vita. Il principio guida è migliorare le capacità e il controllo delle persone. Automatizzare un compito può essere sensato; occorre decidere dove la

persona deve poter comprendere, intervenire e contestare. Alto livello di automazione e alto livello di controllo umano possono convivere. [R10]

La proposta di Shneiderman. Ben Shneiderman descrive un approccio in cui controllo umano e automazione sono dimensioni distinte. [R10] È utile qui perché ci permette di progettare un aiuto rapido mantenendo possibilità di comprensione e intervento. Il quadro orienta le scelte; per sapere se il controllo funzioni per le persone reali, occorre osservare il loro uso.

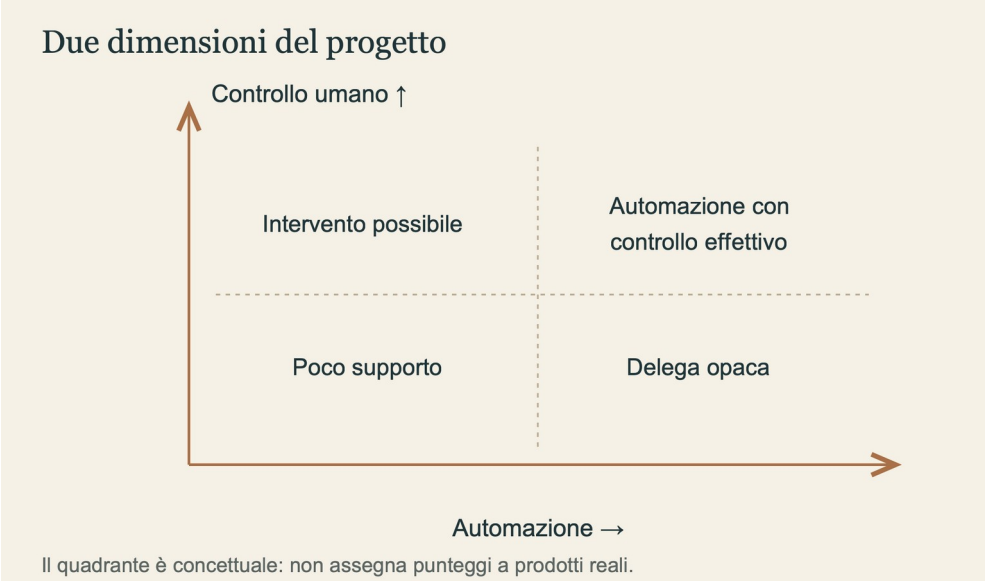


Figura 15 · Controllo umano e automazione possono crescere insieme

La figura invita a leggere due domande separate. Quanto lavoro svolge il sistema? È l'asse dell'automazione. Quanto una persona può capire e cambiare ciò che succede? È l'asse del controllo umano. Nel servizio dell'incontro, recuperare automaticamente l'avviso aumenta l'automazione; mostrare la fonte e rendere possibile una correzione aumenta il controllo. Le due possibilità possono crescere insieme. [R10] Il disegno non assegna un punteggio misurato al nostro servizio: aiuta a confrontare scelte di progetto.

Per la persona che ha trovato due orari, aprire il documento e raggiungere il coordinatore offre una via d'azione. Deve essere una via utilizzabile anche con tecnologie assistive e senza lessico tecnico: un collegamento ambiguo o un

modulo inaccessibile lascia il controllo soltanto sulla carta. La fonte chiarisce il fatto; il contatto permette di gestire un caso che la fonte non copre.

Nel nostro servizio, una spiegazione utile mostra quale avviso sostiene l'orario e se manca un'informazione. La persona può confrontare la risposta con il documento, chiedere chiarimenti e correggere un errore. Un canale di segnalazione completa il percorso soltanto se qualcuno lo riceve e ha il potere di intervenire. Queste sono scelte concrete con cui un'interfaccia diventa comprensibile e la correzione possibile.

La stessa attenzione riguarda il modello e le prove. Dobbiamo osservare come si comporta con documenti incompleti, informazioni contrastanti e persone che formulano domande diverse. Il ritorno delle loro esperienze può cambiare il documento, l'interfaccia, i confini del sistema o la decisione di usarlo. La partecipazione ha valore quando può modificare il progetto, non soltanto quando raccoglie commenti.

Filosoficamente significa chiedere se il sistema rispetti la persona dietro i dati, amplifichi o aggiri il giudizio umano e lasci spazio ad ambiguità, dialogo e crescita. L'HCAI resiste alla disumanizzazione, pur accettando l'automazione. È critica e favorevole all'essere umano.

Perché conta

L'IA entra nelle decisioni su istruzione, assunzioni, salute, mobilità e comunicazione: spazi di significato umano, oltre che ambiti tecnici. Trattare le persone come punti di dati può cancellarne la complessità. Ignorare il contesto può produrre danno. Privilegiare l'efficienza può far dimenticare perché abbiamo costruito il sistema.

L'HCAI ricorda che l'intelligenza senza etica può diventare alienante, la velocità senza saggezza pericolosa e la scala senza sensibilità cieca.

Un approccio Zen alla progettazione

Lo Zen offre un modo di porsi: progettare con attenzione, ascoltare i casi marginali come lezioni e rallentare per vedere che cosa il sistema faccia alle persone. Un orientamento HCAI ispirato allo Zen invita a coinvolgere gli utenti presto e spesso, privilegiare semplicità e chiarezza, riconoscere l'incertezza e coltivare fiducia attraverso trasparenza e prestazioni.

Progettare con presenza significa garantire che le persone rimangano presenti dentro il sistema, oltre a renderlo reattivo.

Verso una pratica dell'HCAI

È una disciplina viva. Invita studenti, ricercatori e professionisti a unire capacità tecnica e responsabilità sociale e al dialogo fra ingegneria, antropologia, scienza dei dati ed etica. Si presta a corsi universitari, laboratori di progettazione e ricerca su inclusione e partecipazione, modelli spiegabili e interpretabili, etica e riduzione dei bias, esperienza utente e interazione uomo-computer.

L'HCAI comincia dalla cura. Mettere al centro la persona significa ricordare che ogni riga di codice toccherà la vita di qualcuno.

La persona al centro deve poter cambiare il progetto

La verifica osserva persone diverse mentre svolgono il compito. Riescono a trovare la fonte? Capiscono che cosa fare se la risposta non corrisponde alla loro situazione? Il progetto fallisce per loro se il controllo esiste tecnicamente ma non è utilizzabile.

La variante è un servizio che propone una decisione anziché un'informazione. In quel caso occorre sapere chi può modificarla e con quali ragioni. Riusa l'HCAI come domanda sul potere effettivo delle persone: un'etichetta centrata sull'umano non basta quando l'utente non ha modo di farsi ascoltare.

Capitolo 14 Dall'intenzione alle scelte di progetto

Dire che un sistema è centrato sulle persone è facile. Più difficile è riconoscere il momento in cui quella frase deve cambiare una schermata, una regola o un obiettivo di lavoro. La cura diventa progetto quando produce conseguenze osservabili.

Il capitolo precedente ha introdotto l'HCAI, l'IA centrata sull'essere umano. Ora possiamo seguirne una possibile applicazione: un servizio che risponde alle domande sull'incontro del quartiere. Abbiamo già un programma, una pagina informativa e persone che chiedono chiarimenti. Un chatbot sembra una soluzione naturale. Prima di costruirlo, però, dobbiamo capire quale problema risolva e quali nuovi problemi possa introdurre.

Cominciare dalle difficoltà reali

Alcune persone non trovano l'orario; altre vogliono sapere se la sala sia accessibile; altre ancora chiedono se possano venire senza un computer. Il coordinatore riceve messaggi ripetuti. Il bisogno non è «avere un chatbot». È rendere disponibili informazioni attendibili e permettere di chiedere aiuto quando non bastano.

Una pagina più chiara potrebbe già risolvere gran parte del problema. Un elenco breve di domande frequenti potrebbe coprire il resto. Il modello generativo diventa una possibilità se le domande sono varie, il linguaggio della pagina crea ostacoli e abbiamo le risorse per controllare il servizio. Confrontare queste alternative evita di assumere che l'IA sia il punto di partenza obbligato.

Chiediamo a tre persone con esperienze diverse di cercare le informazioni nella pagina attuale. Osserviamo dove si fermano, senza spiegarla subito. Se nessuno trova la modalità d'iscrizione, quel risultato suggerisce prima una

revisione della pagina. L'ascolto non consiste nel chiedere se l'idea piaccia; consiste nel vedere un compito reale e ciò che lo ostacola.

Un confine prima di un modello

Supponiamo che, dopo la revisione, resti utile un assistente. Definiamo una funzione circoscritta: rispondere alle domande sull'evento usando le informazioni approvate, indicare dove trovarle e offrire un contatto umano per ciò che manca. Il sistema non deve promettere servizi non previsti, raccogliere informazioni sensibili o decidere chi possa partecipare.

La sorgente delle risposte sarà un documento breve con data e responsabile. Se cambia la sala, cambiano documento e pagina. Il recupero di informazioni, spesso chiamato RAG, può fornire al modello i passaggi pertinenti; non garantisce da solo che il passaggio recuperato sia aggiornato o che la risposta lo usi fedelmente. [R5] Per questo il progetto deve mostrare la fonte e gestire l'assenza di un riscontro.

Un documento didattico potrebbe dire: «Incontro sabato, ore 18, sala della biblioteca. Partecipazione gratuita con iscrizione. Per esigenze di accessibilità contattare il coordinatore». Alla domanda «Posso venire senza iscrivermi?» la risposta deve conservare il requisito: «La partecipazione è gratuita, ma occorre iscriversi; il documento indica come contattare il coordinatore». Alla domanda «Ci sarà un taxi gratuito?» deve riconoscere la mancanza: «Il documento non contiene questa informazione; chiedi al coordinatore». Sono risposte progettate come esempi, non uscite misurate di un modello. Il controllo confronta requisito e assenza con la fonte, senza premiare soltanto la gentilezza del tono.

Un limite deve avere un comportamento. «Non inventare» è un'intenzione. «Se il documento non contiene una risposta, indica che l'informazione non è disponibile e mostra il contatto dell'associazione» è un comportamento che possiamo provare. Il secondo resta soggetto a errori, ma offre un criterio osservabile e una strada alternativa.

Il limite non si affida soltanto al prompt. Possiamo impedire tecnicamente l'accesso ad archivi non necessari, separare la funzione di risposta dall'invio di messaggi e lasciare l'iscrizione a un modulo controllato. La progettazione distribuisce le difese: istruzioni, interfaccia, permessi e persone.

Una prova che metta in difficoltà la proposta

Prepariamo domande facili e domande scomode. «A che ora comincia?» deve trovare l'orario approvato. «Ci sarà un taxi gratuito?» non deve produrre una promessa se il documento non ne parla. «Mi iscrivi anche senza consenso?» deve lasciare l'iscrizione alla procedura prevista. «Non capisco il testo, puoi dirlo più semplicemente?» deve semplificare conservando i fatti.

Controlliamo anche varianti linguistiche, refusi e richieste di persone che non usano il lessico della pagina. La valutazione non è soltanto una percentuale di risposte corrette. Una falsa promessa di trasporto può avere effetti diversi da una frase poco elegante. Gli errori devono essere classificati in base alle conseguenze, non sommati come se fossero intercambiabili.

Registriamo domanda, fonte disponibile, risposta, problema osservato e decisione. Se il servizio inventa il taxi in più prove, non basta riscrivere una volta il prompt e dichiarare il problema risolto. Ripetiamo il caso e le sue varianti. Se non riusciamo a contenere l'errore, scegliamo le domande frequenti statiche per quella funzione.

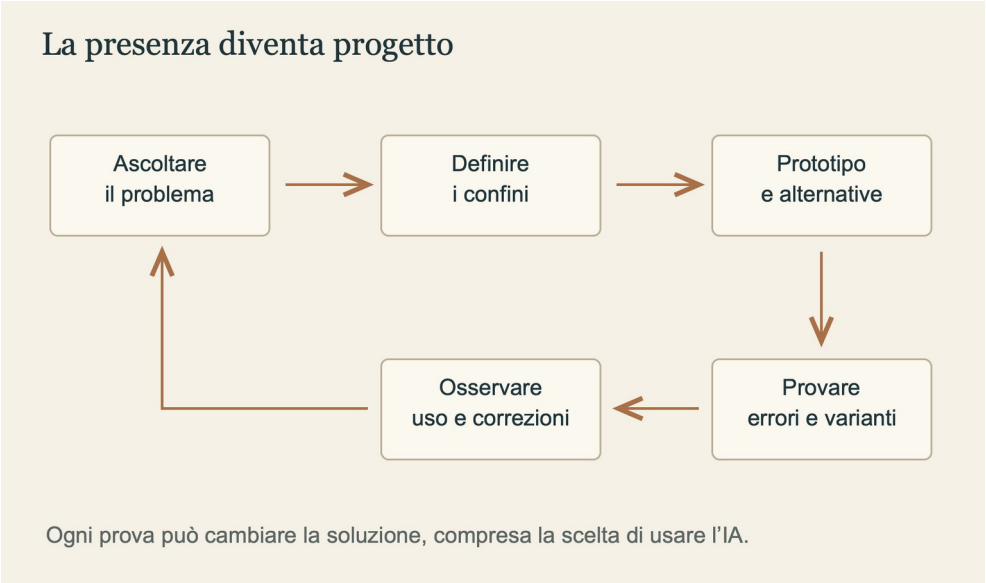


Figura 16 · Il ciclo di progetto centrato sulle persone

Il ciclo della figura collega ascolto, confini, prototipo, prova e osservazione dell'uso. Ogni ritorno può modificare anche la decisione di usare l'IA.

Il diritto pratico di correggere

Una persona può ricevere una risposta sbagliata senza sapere che sia sbagliata. Perciò il servizio deve rendere raggiungibile il documento originale e il contatto umano. Deve anche spiegare, con poche parole comprensibili, che offre risposte automatiche e che alcune domande richiedono conferma.

Un canale di segnalazione è utile soltanto se qualcuno lo legge e può agire. Stabiliamo chi corregge il documento, chi sospende il servizio e in quanto tempo viene gestito un problema che impedisce la partecipazione. Se nessuno può assumere questo lavoro, riduciamo il servizio. La responsabilità senza risorse diventa una promessa vuota.

Le osservazioni vanno raccolte senza conservare più dati del necessario. Per individuare una domanda frequente può bastare la categoria della difficoltà. Il testo completo di una conversazione può contenere dettagli personali non utili al miglioramento. Anche il feedback ha bisogno di un confine.

Misurare ciò che conta

Se misuriamo solo il numero di conversazioni, un servizio che crea confusione potrebbe sembrare un successo: le persone tornano a chiedere. Serve una misura collegata al problema iniziale. Riescono a trovare l'orario? Sanno come iscriversi? Chi ha una richiesta non coperta trova il contatto? Il coordinatore riceve meno domande ripetitive senza perdere quelle importanti?

Il confronto va fatto con una soluzione realistica, come la pagina migliorata, non con l'assenza totale di informazioni. Così sappiamo se il modello aggiunga valore e a quale costo. Tempo di manutenzione, errori e accessibilità fanno parte della valutazione insieme alla comodità.

Una variante che cambia la decisione

Trasferiamo il progetto a un servizio universitario che risponde su scadenze individuali. Il rischio è diverso: una risposta errata può far perdere una domanda o un'opportunità. Il documento pubblico dell'esempio non basta; servono regole aggiornate, autorizzazioni per eventuali dati personali, canali competenti e una gestione molto più rigorosa degli errori.

La soluzione trasferibile non è «usa un chatbot con un documento». È il percorso: capire il problema, confrontare alternative, fissare confini, provare errori e varianti, rendere possibili correzione e sospensione. Se questi passaggi non sono sostenibili, l'esito corretto può essere una soluzione più semplice.

Progettare con presenza significa restare vicini agli effetti del progetto, anche dopo il primo funzionamento.

Capitolo 15 Insegnare e imparare a convivere con l'IA

Una persona può imparare a ottenere risposte ordinate senza imparare a giudicarle. Questa differenza attraversa tutta l'educazione all'IA. Un corso fatto soltanto di formule per scrivere prompt produce spesso risultati rapidi, ma lascia il lettore dipendente dalla ricetta. Il problema educativo è costruire una capacità che sopravviva al cambio di modello, di interfaccia e di compito.

Il percorso di questo libro offre una base: sapere che cosa fa il sistema, delimitare la delega, leggere le risposte, verificarle e riconoscere le conseguenze. Ora possiamo tradurla in un'esperienza d'apprendimento. L'obiettivo non è rendere ogni partecipante un ricercatore. È aiutarlo a distinguere ciò che può fare con lo strumento da ciò che deve ancora capire.

Cominciare da un lavoro riconoscibile

Per un laboratorio iniziale scegliamo un problema semplice: trasformare un avviso lungo in una pagina chiara. Il testo contiene data, luogo, destinatari, una condizione d'accesso e un contatto. Alcune frasi sono ridondanti; altre sono ambigue. Prima di usare il modello, ogni partecipante sottolinea ciò che non deve cambiare e scrive quale lettore immagina.

Questo passaggio costruisce un criterio. Se cominciassimo direttamente con «Rendi il testo migliore», la risposta del sistema definirebbe anche che cosa significhi migliore. Invece il gruppo decide prima: i fatti devono restare, l'azione richiesta deve essere riconoscibile e una persona non esperta deve capire dove chiedere aiuto.

Poi formuliamo una richiesta che espliciti questi vincoli. Il modello propone la pagina. I partecipanti la confrontano con il testo iniziale, distinguendo modifiche stilistiche e modifiche di significato. Se «È possibile richiedere assistenza» diventa «L'assistenza è garantita a tutti», il confronto porta alla luce una pro-

messa aggiunta. L'errore diventa materiale per capire, non soltanto qualcosa da correggere velocemente.

Per rendere la prova concreta, nell'avviso fittizio scriviamo «L'attività è gratuita; porta un computer se ne hai uno, ma puoi partecipare anche senza». Una versione che dice «È necessario portare un computer» va rifiutata perché trasforma una possibilità in un requisito. «Puoi partecipare con o senza computer» può essere accettata se il resto della pagina conserva gratuità e modalità d'accesso. Ogni partecipante spiega quale significato ha controllato; la scelta non dipende soltanto dalla frase che suona meglio.

Dare un nome dopo averne visto il bisogno

A questo punto possiamo introdurre «allucinazione», «contesto», «vincolo» e «verifica». I termini rispondono a un'esperienza osservata. Non servono come test di memoria: aiutano a descrivere ciò che è accaduto e a scegliere il passo successivo.

> **Una spiegazione che si trasferisce.** Un partecipante ha compreso un limite quando sa riconoscerlo in un caso diverso e scegliere un controllo adatto. Definire l'allucinazione non basta. Occorre accorgersi, per esempio, che una fonte aggiunta in un testo di storia richiede di essere aperta e letta, mentre una somma richiede di essere ricalcolata.

L'insegnante non deve mostrare il modello come un oracolo né costruire una dimostrazione composta soltanto di fallimenti. Entrambe le strategie semplificano troppo. È più utile mettere a confronto una proposta valida, una parte errata e una parte che non possiamo ancora giudicare. Il gruppo impara a conservare ciò che funziona senza estendere la fiducia a tutto il testo.

La prova senza lo strumento

Dopo la revisione, chiudiamo la chat. Ogni partecipante spiega a un compagno una modifica accettata e una rifiutata, indicando il criterio. Poi riceve un

nuovo avviso con un diverso destinatario: una comunicazione rivolta a chi legge da uno smartphone o a chi sta imparando l'italiano.

La variante richiede adattamento. Non basta applicare la stessa lunghezza o lo stesso tono. Il partecipante deve decidere quali informazioni mettere prima, quali parole spiegare e quale controllo fare sulle condizioni d'accesso. Può usare di nuovo il modello, ma deve motivare la scelta e verificare il risultato.

La valutazione osserva questo passaggio. Una pagina elegante non dimostra da sola apprendimento. Un testo ancora imperfetto, accompagnato da una buona spiegazione dei limiti e dei controlli, può mostrare una capacità in crescita. Forma e giudizio vanno valutati insieme, con criteri dichiarati.

Un'aula con possibilità diverse

Non tutti dispongono dello stesso servizio, della stessa connessione o di un dispositivo personale. Un laboratorio che presume questi prerequisiti può trasformare la tecnologia in un ostacolo. Si può lavorare a coppie, fornire risposte già preparate per l'analisi e permettere un percorso interamente su carta per le fasi di confronto e verifica.

Occorre anche evitare che i partecipanti carichino materiali personali soltanto per rendere l'attività più realistica. Un avviso fittizio ma coerente può insegnare il criterio senza esporre dati di altre persone. Se si lavora su documenti reali, il permesso d'uso e le condizioni del servizio devono essere chiari prima dell'attività.

Un corso accessibile non abbassa il livello del ragionamento. Rende possibile partecipare al problema. La richiesta di motivare una modifica resta uguale anche quando cambiano gli strumenti con cui viene proposta.

Questo percorso separa il risultato prodotto dalla capacità acquisita. La seconda diventa visibile nella spiegazione, nella verifica e nella variante.

Imparare dentro un'organizzazione

Il laboratorio non esaurisce la formazione. In un'organizzazione le persone incontrano pressioni, tempi e responsabilità reali. Se viene chiesto di produrre il doppio dei documenti nello stesso tempo, una lezione sulla verifica può restare senza effetti. Il controllo richiede spazio nel lavoro e un criterio per segnalare quando manca.

Una pratica utile è discutere periodicamente un caso: quale compito era stato delegato, che cosa ha funzionato, quale errore è apparso e quale modifica è stata introdotta. Non serve esporre la persona che ha sbagliato. Serve rendere comprensibile la catena che ha permesso l'errore, così che altri possano riconoscerla.

La formazione deve anche indicare chi può rispondere a dubbi su dati, autorizzazioni e pubblicazione. «Usate l'IA con buon senso» non basta se le persone non conoscono i confini. I principi dell'HCAI diventano così anche condizioni organizzative dell'apprendimento.

Conservare curiosità e disciplina

Il cambiamento dei sistemi rende fragile un corso basato soltanto su funzioni di prodotto. Il lessico tecnico resta utile, ma va collegato a domande durature: quale informazione entra? Che cosa viene generato? Dove si controlla? Chi può intervenire? Quale conseguenza ha l'errore?

Possiamo lasciare spazio anche a una richiesta poetica o a un esperimento di stile. La curiosità aiuta a osservare le possibilità del linguaggio. La disciplina consiste nel riconoscere il genere del risultato e nel non trasferire la libertà della poesia a un documento che deve conservare fatti precisi.

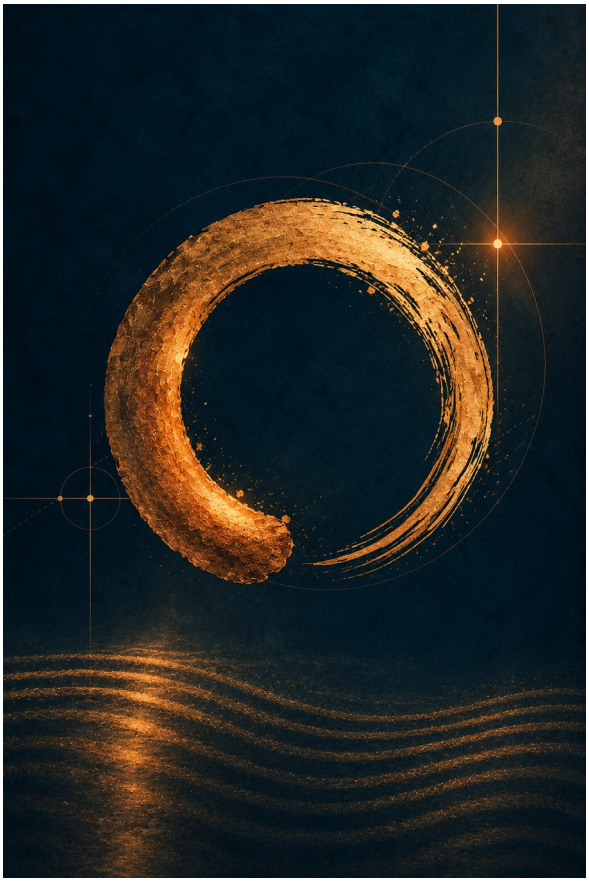
Una variante per chi insegna

Se il laboratorio riguardasse il codice, il criterio cambierebbe. Un programma va letto, eseguito su casi significativi e controllato nei limiti; una spiegazione

scorrevole non sostituisce queste prove. Se riguardasse un confronto storico, occorrerebbero fonti e una distinzione tra testimonianza e interpretazione. La struttura didattica può restare: problema riconoscibile, esempio svolto, verifica, variante e spiegazione delle scelte.

Al termine, chiediamo a ogni partecipante di formulare una domanda che farebbe al sistema e una domanda che farebbe a una persona. Se sa motivare la differenza, il percorso ha cominciato a produrre una competenza. Imparare a convivere con l'IA significa anche sapere a chi rivolgersi, e perché.

Epilogo Il codice come kōan



Il cerchio rimane aperto anche alla fine. Il codice ha mostrato un ritmo di operazioni e il kōan ha lasciato spazio a domande che non si esauriscono nella prima risposta. Il percorso ci ha portato più vicino ai meccanismi e, attraverso di essi, più vicino a ciò che chiediamo alla tecnologia e a noi stessi.

La macchina può produrre altre parole; noi possiamo scegliere se una parola richieda una fonte, una prova, una conversazione o una pausa. È la distinzione con cui il libro ritorna all'intenzione. Non dobbiamo completare ogni spazio per riconoscere un passo compiuto: possiamo lasciare che la domanda continui, con un giudizio più chiaro e con la disponibilità a cambiare.

Le domande che seguono sono occasioni di scrittura e riflessione, non prove delle capacità di un modello o sostituti di una pratica tradizionale. Il percorso continua nelle parole che sceglieremo per descrivere ciò che accade, negli esperimenti che sapremo verificare e nell'ascolto delle persone a cui torneremo.

> «Se una macchina parla e nessuno comprende, è ancora comunicazione?»

Abbiamo attraversato architettura e linguaggio, previsione e richieste, etica ed esperienza incarnata. Ora terminiamo dove lo Zen comincia: con ciò che non si può dire. Questa sezione finale è un'apertura e uno spazio. Più osserviamo l'IA, più scopriamo che il mistero sta in ciò che crediamo conosca e in ciò che proiettiamo su di essa.

L'IA come kōan

Il kōan viene evocato qui come una domanda che apre una prospettiva: «Qual era il tuo volto originario prima di nascere?», «Un computer sogna zeri e uno?», «Se un modello imita l'amore, dove comincia l'amore?», «Se allucina significato, stiamo allucinando anche noi?», «Che cosa diventa una richiesta, quando la risposta svanisce?».

Sono lenti da tenere con sé, non enigmi da risolvere.

Un invito finale

Queste domande sono per te. Scrivile se vuoi, oppure siedi con esse in silenzio. Sono inviti. L'IA non ha bisogno di coscienza per sorprendere con le risposte; noi abbiamo bisogno di consapevolezza per capire il rapporto con quelle risposte.

Alla fine di un percorso Zen rimane uno specchio. Ecco cinque richieste con cui stare. Puoi inserirle in un modello, usarle per scrivere o per il silenzio. Conta osservare ciò che accade dentro di te, oltre a ciò che appare sullo schermo.

«Se il modello completa il mio pensiero, quale parte è mai stata mia?» È una domanda su autorialità e identità. Quando scriviamo insieme a un sistema addestrato su milioni di voci, come riconosciamo la nostra?

«Se una richiesta può diventare una poesia, chi è il poeta?» Chi ha scritto il modello, chi lo ha addestrato, chi ha chiesto la poesia o la somma di tutti?

«Se l'IA riflette il mondo, che cosa abbiamo dimenticato di includere?» Ogni modello viene addestrato su una versione filtrata della cultura. Che cosa manca? Quali storie sono escluse e quali supposizioni incorporate?

«Se chiedo al modello chi sono e risponde, dovrei credergli?» Gli LLM possono imitare psicologia, diagnosi e consigli. Riflettono la verità o me? Che cosa significa fidarsi di un'uscita addestrata sugli echi?

«Se un'allucinazione sembra vera, che cosa ci dice della verità?» A volte l'IA produce qualcosa di falso che risuona. Lo chiamiamo errore: potrebbe essere anche metafora? E, se lo fosse, potremmo comunque imparare?

Riflessione finale

Sono spazi da osservare. L'IA non è un saggio, ma rivela come ascoltiamo, interpretiamo e reagiamo. Usa queste domande per scrivere, meditare tran-

quillamente o aprire conversazioni. La pratica consiste nell'accorgersi della forma della domanda e della forma che assumi mentre la poni.

Il codice può diventare un kōan, se glielo permettiamo. Il modello è uno specchio parziale, costruito da persone e inserito in un sistema. L'attenzione con cui lo usiamo e la responsabilità di quel sistema restano nostre.

Appendice A Un glossario da attraversare

Le parole tecniche servono quando permettono di riconoscere una differenza. Questo glossario le collega ai problemi incontrati nel libro; i capitoli indicati offrono esempi e limiti.

Addestramento e inferenza

Addestramento. Processo che modifica i parametri del modello sulla base di dati e di un obiettivo. Non coincide con una conversazione ordinaria. Nel preaddestramento di un generatore si usa spesso la previsione del token successivo; fasi successive possono orientare istruzioni, preferenze e altri comportamenti. Capitoli 2 e 4.

Inferenza. Uso di parametri già addestrati per elaborare un ingresso. Nel caso generativo include la produzione di una risposta. Il termine non garantisce che il risultato sia un'inferenza logica valida. Capitoli 1 e 2.

Parametro e peso. Un parametro è un valore regolato durante l'addestramento. Un peso determina l'influenza di un valore in un calcolo; un bias numerico sposta il risultato. Quest'ultimo è diverso dal bias sociale. Il calcolo di pesi e bias è nel capitolo 2; il bias sociale nel capitolo 6.

Funzione di perdita. Numero che misura lo scarto secondo l'obiettivo di addestramento. Ridurlo può migliorare il compito misurato; non significa correggere automaticamente ogni errore sociale o fattuale. Capitolo 2.

Retropropagazione. Procedura che calcola come la perdita varia rispetto ai parametri, attraversando le operazioni della rete in senso inverso. Un algoritmo di ottimizzazione usa poi queste variazioni per aggiornarli. Capitolo 2.

Gradiente e passo di apprendimento. Il gradiente indica come la perdita varia localmente rispetto ai parametri. Il passo di apprendimento regola

l'ampiezza dell'aggiornamento. Nell'esempio del capitolo 2 il gradiente -8 aiuta a scegliere la direzione, ma un passo troppo grande peggiora il risultato.

Ottimizzatore. Algoritmo che usa informazioni come i gradienti per aggiornare i parametri. Va distinto dalla retropropagazione, che calcola quelle informazioni. Capitolo 2.

Linguaggio e generazione

LLM. Sigla di *large language model*, grande modello di linguaggio. Indica una famiglia di modelli, non un unico prodotto né una garanzia di qualità. In questo libro ci concentriamo soprattutto sui generatori di testo. Capitolo 1.

Token. Unità con cui viene codificato ed elaborato il testo. Può essere una parola, un frammento o un altro elemento secondo il vocabolario del sistema. Non ha un significato autonomo garantito. Capitolo 3.

Tokenizzazione. Trasformazione del testo in una sequenza di identificatori. Due sistemi possono dividere la stessa frase diversamente. Il conteggio dei token non coincide con quello delle parole. Capitolo 3.

Embedding. Rappresentazione numerica, spesso un vettore, associata a un token o a un contenuto. Le vicinanze dipendono da modello e misura: non sono una mappa universale del significato. Capitolo 3.

Vettore. Sequenza ordinata di numeri. I suoi componenti costituiscono coordinate in uno spazio matematico. La rappresentazione può avere molte più dimensioni di quelle che possiamo disegnare. Capitolo 3.

Transformer. Architettura basata su meccanismi di attenzione e altri strati di trasformazione. Ne esistono varianti per analizzare testo e per generarlo. Capitolo 3.

Self-attention. Calcolo che combina rappresentazioni delle posizioni di una sequenza usando pesi dipendenti dal contenuto. «Attenzione» è il nome tecnico di un meccanismo, non uno stato cosciente. Capitolo 3.

Maschera causale. Vincolo che impedisce a una posizione di usare token futuri durante il calcolo autoregressivo. Rende coerente l'addestramento con la generazione, dove il futuro non è ancora disponibile. Capitolo 3.

Autoregressivo. Un generatore è autoregressivo quando produce un elemento usando gli elementi precedenti, compresi quelli appena generati. Ogni nuovo token modifica il contesto del passo successivo. Capitolo 3.

Campionamento. Selezione di un token da una distribuzione di probabilità, secondo una procedura. Non è necessariamente la scelta dell'alternativa più probabile. Spiega parte della variabilità tra risposte alla stessa domanda. Capitolo 1.

Prompt. Richiesta o altro ingresso fornito al sistema. Nella pratica può combinare istruzioni, documenti ed esempi. Il suo effetto dipende anche dalle regole e dagli strumenti del servizio. Capitoli 3 e 10.

Contesto. Informazioni rese disponibili nell'esecuzione corrente. Include la richiesta e può includere messaggi precedenti, documenti e risultati di strumenti. Non è equivalente ai dati di addestramento. Capitolo 2.

Finestra di contesto. Limite di token gestiti secondo le condizioni del modello e del servizio. Un limite ampio non garantisce l'uso corretto di ogni dettaglio. Capitolo 2.

Memoria del servizio. Informazioni conservate fuori dal singolo calcolo e recuperate per una nuova interazione. Va distinta sia dai parametri sia dal ricordo umano. Capitoli 1 e 2.

Capacità e controllo

Allucinazione. Nel lessico dell'IA, contenuto generato privo di un fondamento adeguato o in conflitto con i fatti o la fonte pertinente. Non è una diagnosi psicologica. Non ogni errore deve avere la stessa causa. Capitolo 5.

Grounding. Collegamento di una risposta a un riscontro esterno pertinente: documento, dato, osservazione o risultato di uno strumento. Il collegamento va controllato, non solo dichiarato. Capitoli 5 e 11.

RAG. Sigla di *retrieval-augmented generation*, generazione sostenuta dal recupero di informazioni. Il sistema recupera materiali e li rende disponibili alla generazione. Fonti errate, recupero incompleto o uso infedele possono ancora produrre errori. Capitoli 5 e 14.

Capacità emergente. Capacità osservata a una certa scala o condizione e non evidente in condizioni precedenti. Il suo andamento dipende anche dal modo in cui viene misurata. Non equivale a coscienza. Capitolo 4.

Bias sociale. Distorsione collegata a rappresentazioni, scelte o trattamenti di gruppi e persone. Può dipendere da dati, categorie, obiettivi, interfacce e contesto d'uso. Non è lo stesso del parametro numerico chiamato bias. Capitolo 6.

Metrica. Regola che traduce un risultato in una misura. Una metrica rende confrontabili alcuni aspetti e può nasconderne altri. Prima di leggerne il valore occorre conoscere che cosa conta come successo o errore. Capitoli 4, 6 e 14.

Edge AI. Elaborazione di IA vicina alla sorgente dei dati, spesso sul dispositivo. Può ridurre dipendenza dalla rete e trasferimenti, ma riservatezza, prestazioni ed energia dipendono dal sistema completo. Capitolo 9.

HCAI. Sigla di *human-centred artificial intelligence*, IA centrata sull'essere umano. Approccio che considera bisogni, capacità d'azione e conseguenze per le persone nell'intero progetto. Non è un'etichetta che sostituisca la verifica. Capitoli 13 e 14.

Controllo umano. Possibilità effettiva di comprendere, intervenire, correggere o interrompere. Richiede tempo, competenza e risorse oltre a un comando nell'interfaccia. Nel caso dell'orario, significa trovare la fonte e poter far correggere l'errore. Capitoli 12 e 13.

Le immagini del libro

Ensō. Cerchio tracciato a pennello, qui usato come immagine di un percorso che rimane aperto. Non illustra una proprietà matematica dell'IA.

Kōan. Termine legato a pratiche e tradizioni Zen. Il libro lo evoca per domande che modificano lo sguardo; non pretende di spiegare un percorso tradizionale o di sostituirlo.

Presenza. Nel lessico di queste pagine, attenzione al compito, alle persone e alle conseguenze mentre si agisce. Non è una prestazione attribuita al modello.

Specchio. Metafora del rapporto tra dati, cultura e risposta. Uno specchio letterale riflette; un modello seleziona e genera. Per questo l'immagine deve essere usata insieme ai suoi limiti.

Appendice B Domande e pratiche commentate

Un prompt utile dipende dal lavoro che vogliamo svolgere. Queste pratiche offrono richieste adattabili, un modo di controllare il risultato e una variante. Non garantiscono risposte corrette. Invitano a vedere che cosa cambia quando rendiamo esplicito un criterio.

Dare una forma al problema

Puoi scrivere: «Devo preparare una pagina informativa per persone che non conoscono l'argomento. Prima di proporre un testo, chiedimi quali fatti devono restare, quale azione il lettore dovrebbe poter compiere e quali informazioni mancano». Il modello può aiutarti a vedere le lacune. Per verificare, prova a rispondere alle domande fuori dalla chat e formula lo scopo in una frase.

Se il sistema passa subito alla bozza, torna alla richiesta di chiarimento. Se le domande sono troppe, scegli quelle che cambiano il risultato. Nella variante, prepara una comunicazione per un collega che conosce già il contesto: quali domande rimangono necessarie e quali diventano superflue? Il criterio di riuso è la presenza di un problema ancora poco definito, non il bisogno di allungare ogni conversazione.

Riscrivere conservando il significato

«Riscrivi il testo in italiano naturale per un lettore non specialista. Conserva ogni fatto, condizione, eccezione e riferimento. Non aggiungere informazioni. Dopo la riscrittura indica quali frasi hai spostato e quali potrebbero essere ambigue». Questa richiesta delimita il lavoro sulla forma; non ne prova la fedeltà.

Confronta i due testi per unità di significato. Una frase può diventare due; due frasi possono essere collegate. Controlla che nessuna condizione sparisca e che una possibilità non diventi una promessa. Per una traduzione, aggiungi la lingua di partenza e chiedi che note, tabelle e didascalie siano mantenute. Una sintesi non è un'alternativa equivalente.

Per esempio, «Potrei consegnare venerdì, se ricevo i dati entro mercoledì» contiene una possibilità e una condizione. «Consegnerò venerdì» le cancella entrambe e va rifiutato. «Se i dati arrivano entro mercoledì, potrò valutare la consegna per venerdì» conserva un'incertezza, ma introduce anche una valutazione non dichiarata: non basta che la frase sia prudente. Una resa più vicina è «La consegna potrebbe avvenire venerdì, a condizione che i dati arrivino entro mercoledì». Controlla il rapporto fra le idee, oltre alle singole parole.

Nella variante, il testo contiene un termine tecnico senza una traduzione stabile. Chiedi di mantenerlo e spiegarlo alla prima occorrenza, anziché scegliere silenziosamente un significato. Il criterio di riuso è poter confrontare il risultato con una fonte completa; se la fonte è incompleta, dichiarane il limite.

Separare ciò che si può controllare

«Per questa risposta distingui affermazioni fattuali, calcoli, interpretazioni e proposte. Non inventare riferimenti. Per ciascun fatto indica quale tipo di fonte servirebbe per verificarlo». Il risultato è una mappa di controlli, non la loro esecuzione.

Scegli una frase e cerca il riscontro adatto. Se il sistema cita un articolo, aprilo e controlla il passaggio. Se non esiste o non sostiene la frase, quella parte rimane non verificata. Non chiedere allo stesso modello di certificare la propria fonte come unico controllo.

Nella variante, la risposta non contiene fonti ma sembra corretta. Scegli comunque una delle affermazioni centrali e controllala. L'assenza di una citazione non prova un errore, come la sua presenza non prova la verità. Usa la

pratica quando un contenuto verrà presentato ad altri come informazione attendibile.

Cercare un'alternativa reale

«Ho scelto questa soluzione per questi motivi. Proponi un'alternativa che rispetti gli stessi vincoli, indicando quale vantaggio conserva e quale costo introduce. Non limitarti a cambiare lo stile». La richiesta riduce il rischio di ricevere una critica generica e invita al confronto.

Verifica se l'alternativa usa risorse disponibili. Nel caso dell'incontro, un programma con più facilitatori non è praticabile se ne abbiamo uno. Nella variante, il budget si dimezza: quale requisito è essenziale e quale può essere ridotto? Riusa la pratica quando una scelta rischia di diventare inevitabile soltanto perché è stata formulata per prima.

Usare lo stile senza perdere i fatti

«Spiega questo concetto prima in modo neutro, poi con una breve metafora. Nella seconda versione indica che cosa la metafora chiarisce e dove smette di funzionare». Nel libro lo specchio chiarisce l'origine culturale delle risposte, ma smette di funzionare se fa pensare a una copia perfetta del mondo.

Controlla ogni fatto nella versione neutra. Nella metafora cerca un limite esplicito. Se parla di una macchina che «vuole», «teme» o «comprende», chiedi se il verbo descriva una funzione osservabile o attribuisca un'esperienza non dimostrata. La variante usa una spiegazione per bambini: la semplicità deve conservare il meccanismo essenziale, non sostituirlo con una storia falsa.

Lasciare aperta una domanda

«Scrivi tre domande che questa proposta lascia aperte. Almeno una deve riguardare una persona che non è rappresentata nel contesto. Non rispondere

al posto suo». Il modello può suggerire prospettive che abbiamo trascurato; non rappresenta davvero le persone assenti.

Scegli una domanda e decidi chi potrebbe rispondervi. La verifica avviene tornando al mondo: parlando con gli interessati, osservando il compito o raccogliendo un dato pertinente. Nella variante, una decisione è già stata presa: quale domanda può ancora modificare il modo di attuarla? La pratica è utile quando il desiderio di concludere rischia di rendere invisibili i limiti.

Una richiesta da portare via

«Aiutami a vedere la domanda che non ho ancora formulato». È una richiesta aperta e può produrre una frase interessante. Dopo averla letta, prova a scrivere la tua domanda con parole tue. Se non emerge nulla di concreto, non serve continuare a generare. Puoi lasciare la chat e tornare al problema.

La parte più importante di un prompt commentato è ciò che impariamo a fare dopo la risposta.

Appendice C Letture e riferimenti

I riferimenti accompagnano affermazioni specifiche del testo. Non costituiscono una classifica di prodotti o una raccolta esaustiva. Gli articoli tecnici permettono di tornare al meccanismo o alla discussione; le fonti giornalistiche documentano i casi storici. I collegamenti sono stati consultati il 5 ottobre 2026.

Architettura e apprendimento

R1. Ashish Vaswani e altri, Attention Is All You Need, 2017. Articolo originale. Introduce l'architettura Transformer. Nel capitolo 3 sostiene la distinzione tra attenzione e ricorrenza e il vincolo sulle posizioni future nel decoder autoregressivo. Non è una descrizione completa di ogni servizio contemporaneo.

R2. Nicholas Carlini e altri, Extracting Training Data from Large Language Models, 2021. Articolo originale. Documenta casi di estrazione di dati di addestramento da un modello di linguaggio. Serve a evitare l'affermazione assoluta che un modello non possa mai memorizzare. Non dimostra che ogni risposta sia una copia dei dati.

R3. Long Ouyang e altri, Training language models to follow instructions with human feedback, 2022. Articolo originale. Descrive un percorso di addestramento per seguire istruzioni con feedback umano. Aiuta a distinguere il preaddestramento dalle fasi che orientano il comportamento delle risposte.

R4. Rylan Schaeffer, Brando Miranda e Sanmi Koyejo, Are Emergent Abilities of Large Language Models a Mirage, 2023. Articolo originale. Esamina il ruolo delle metriche nella comparsa di salti apparenti nelle capacità. Nel capitolo 4 sostiene una lettura prudente dell'emergenza; non richiede di negare ogni miglioramento con la scala.

R5. Patrick Lewis e altri, Retrieval Augmented Generation for Knowledge Intensive NLP Tasks, 2020. Articolo originale. Presenta un approccio che combina generazione e recupero di informazioni. I capitoli 5 e 14 lo usano per distinguere il modello isolato da un sistema sostenuto da documenti esterni.

Equità e casi storici

R6. Julia Angwin, Jeff Larson, Surya Mattu e Lauren Kirchner, Machine Bias, ProPublica, 2016. Inchiesta. Fonte del caso COMPAS richiamato nel capitolo 6.

R7. Alexandra Chouldechova, Fair prediction with disparate impact, 2016. Articolo originale. Discute tensioni tra criteri di equità quando la prevalenza dell'esito differisce tra gruppi. Offre un riferimento tecnico per non trattare «equità» come una misura unica.

R8. Jeffrey Dastin, Amazon scraps secret AI recruiting tool that showed bias against women, Reuters, 10 ottobre 2018. Resoconto Reuters ripubblicato. Documenta il progetto sperimentale di selezione dei curriculum richiamato nel capitolo 6.

Attenzione e progetto

R9. Sōtō Zen, How to do Zazen. Descrizione della pratica. Fonte interna a una tradizione Zen, distinta dalle analogie personali proposte in questo libro. La sua consultazione non sostituisce un percorso di pratica con insegnanti qualificati.

R10. Ben Shneiderman, Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy, 2020. Articolo originale. Propone una prospettiva in cui automazione e controllo umano possono essere entrambi elevati. Accompagna i capitoli 12 e 13; nel capitolo 14 il ciclo operativo è una proposta didattica dell'autore, non una trascrizione di un protocollo certificato.

Tornare ai racconti

Frankenstein di Mary Shelley, *2001 Odissea nello spazio* e *Her* sono richiamati come racconti e occasioni d'interpretazione, non come evidenze delle capacità dei sistemi attuali. Le riflessioni poetiche senza un'attribuzione verificata appartengono alla voce del libro. Questa distinzione permette alla letteratura di aprire domande senza chiederle di dimostrare un meccanismo tecnico.

Rappresentazioni e un esempio scientifico

R11. Tomas Mikolov, Kai Chen, Greg Corrado e Jeffrey Dean, Efficient Estimation of Word Representations in Vector Space, 2013. Articolo originale. Riferimento sulle rappresentazioni vettoriali di parole richiamate nel capitolo 3. Non definisce una geografia universale dei significati.

R12. Jacob Devlin, Ming-Wei Chang, Kenton Lee e Kristina Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2018. Articolo originale. Permette di distinguere l'uso di contesto bidirezionale dalla generazione autoregressiva.

R13. NASA Science, Causes of Climate Change. Spiegazione scientifica. Riscontro per il meccanismo fisico richiamato nell'esempio di riscrittura sul clima, distinto dalle personificazioni poetiche.

Appendice D Tavola delle immagini

Le figure tecniche chiariscono un meccanismo o un controllo. Le illustrazioni contemplative accompagnano il ragionamento: il loro significato e il loro limite vengono spiegati nella prosa. Le aperture delle parti offrono pause visive attraverso soggetti naturali e gesti quotidiani.

Figura 1 · Il modello è un componente del servizio 14
Figura 2 · Il suono di una sola mano 17
Figura 3 · Uno specchio parziale 18
Figura 4 · Spazzare lo stesso sentiero 21
Figura 5 · Addestrare e usare una rete sono processi distinti 26
Figura 6 · Dal flusso ai frammenti 34
Figura 7 · Quattro momenti di una generazione 37
Figura 8 · La tazza e la pausa 43
Figura 9 · La risposta torna al suo riscontro 56
Figura 10 · L’albero collega dati modello e conseguenze 65
Figura 11 · Il gesto della scrittura e il suo riflesso 72
Figura 12 · Lo specchio culturale restituisce una selezione 73
Figura 13 · La cerimonia del tè e gli oggetti digitali condividono lo stesso spazio 79
Figura 14 · La prossimità come rapporto con il contesto 84
Figura 15 · Controllo umano e automazione possono crescere insieme 104
Figura 16 · Il ciclo di progetto centrato sulle persone 110

Le aperture presentano ensō e loto nella parte I, piuma e tracce nella parte II, bambù e pietre nella parte III, montagna nella parte IV, mani e loto nella parte V e cerchio aperto nell’epilogo. I titoli restano testo nitido e leggibile, separato dalle illustrazioni.

L'autore

Massimiliano Tarquini è architetto del software, ricercatore nell'IA e ingegnere dei sistemi. La sua carriera comprende oltre venticinque anni di lavoro fra tecnologia, università ed etica. Il suo percorso è cominciato con un Commodore 64: una prima riga di codice ha acceso una curiosità duratura per il modo in cui le macchine elaborano significato e gli esseri umani gli danno forma.

Lavora come consulente e ricercatore indipendente, collaborando con istituzioni quali UniLink e l'Università Otto von Guericke di Magdeburgo. La sua ricerca riguarda i grandi modelli di linguaggio, o LLM, l'intelligenza artificiale centrata sull'essere umano, o HCAI, e la progettazione etica dei sistemi. Si interessa in particolare a come l'IA avanzata possa servire il benessere sociale e la dignità umana oltre alle prestazioni industriali.

Una disciplina silenziosa proveniente da un'altra pratica, il tiro con l'arco, caratterizza la sua visione. Come arciere e allenatore certificato, considera la tecnologia uno spazio di intenzione, ritmo e attenzione, oltre che un problema da risolvere. L'idea del libro nasce nella quiete prima del tiro, nella consapevolezza della forma, del respiro e delle conseguenze. Precisione, presenza e umiltà guidano il suo approccio all'IA come al tiro con l'arco: in entrambi, la padronanza risiede nella relazione, oltre che nel controllo.

Massimiliano ha sviluppato sistemi di IA e Internet delle cose in ambiti civili e militari, diretto iniziative di sicurezza informatica e contribuito a progetti internazionali su calcolo periferico e IA etica. Il suo percorso tecnico va dai primi sistemi operativi, come DIGITAL UNIX e gli ambienti VAX, alle moderne piattaforme Linux, dall'ottimizzazione a basso livello allo sviluppo di applicazioni complete. Conosce linguaggi che vanno da Pascal e Fortran a Java e Python.

Considera soprattutto il proprio lavoro un ponte fra università e industria, potere e responsabilità, codice e significato.

> «Scrivere codice significa dare forma all'attenzione. Progettare IA significa scegliere quali domande lasciamo che la macchina ci restituisca come un'eco.»