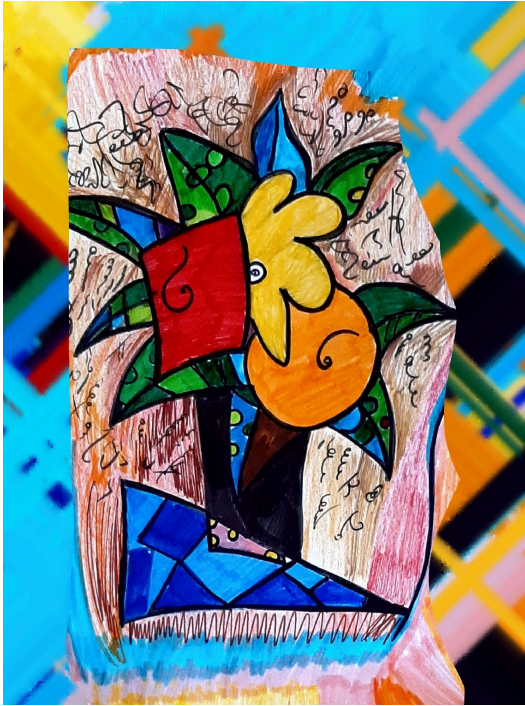


Temi in classe



LETTURE 279

di Stefano Penge

Ma allora cosa si può fare con questa benedetta intelligenza artificiale generativa nell'educazione? Va scansata, ignorata altezzosamente?

Certo che no. Perché al di là della sua forma attuale (il prompt appunto) si tratta di un passaggio importante non solo tecnico, ma culturale e sociale. La frenesia con cui persone che hanno ignorato fino a ieri l'esistenza stessa di software "intelligenti" si sono buttate a creare testi e immagini sulla base di ricette ha una radice che merita di essere studiata e capita.

Si potrebbero "importare" alcuni dei temi toccati fin qui all'interno delle materie e delle lezioni tradizionali. Le storie (della filosofia, della letteratura, dell'arte) sono discipline in cui è facile trovar loro posto.

Per esempio:

- i prodotti dell'IA generativa sono analitici o sintetici? A priori o a posteriori?
- il symbolum niceno-costantinopolitano e la differenza tra generazione e creazione;
- copiare è bello: novità e tradizione tra classicismo e romanticismo;
- automi nella letteratura, da Esiodo a Čapek;
- Narciso e l'amore per la propria immagine, da Caravaggio a Freud;
- ...

In sostanza si tratterebbe di ricondurre la novità nell'alveo della tradizione. Un procedimento tipico della scuola, che tranquillizza tutti, ma a mio avviso rischia di non porta a grandi passi avanti.

Forse si può fare qualcosa di più.

Intanto la discussione che precede e segue l'uso del prompt può più essere utile e interessante. Probabilmente questa è la parte più importante di ogni utilizzo didattico di tecnologie non create esplicitamente per l'apprendimento, dai laboratori linguistici alle LIM: ricostruirne la genesi e gli obbiettivi, provare a capirne il funzionamento, valutarne il bilancio guadagni-concessioni.

Una questione centrale per la generazione dei testi è quella del feedback. Le tecnologie usano spesso il feedback, anzi si può dire che la cibernetica (la "nonna" dell'intelligenza artificiale) sia nata proprio da quest'idea di retroazione. Un esempio che conosciamo tutti: il riscaldamento gestito tramite un termostato. Anche i motori di ricerca usano i feedback degli utenti, ma in maniera trasparente: se un utente clicca su uno dei risultati della lista, il motore suppone che quel risultato sia correlato altamente alla richiesta. Questo ne aumenta la probabilità di essere nuovamente presentato ai primi posti come risultato ad una ricerca uguale o simile.

L'uso di un motore di ricerca modifica i risultati delle prossime ricerche, non solo di quell'utente ma di tutti. E' un po' come attivare il servizio che riceve di dati dai satelliti GPS e usarlo per le mappe di Google: non ci si limita a ricevere un servizio, ma si *contribuisce* a costruirne il valore [<https://www.stefanopenge.it/wp/mappe-personali/>].

Con il paradosso noto per cui se migliaia di automobilisti, per evitare un ingorgo sull'autostrada segnalato dal navigatore, prendono l'uscita prima dell'ingorgo, quell'uscita dopo pochi minuti si intasa peggio dell'autostrada stessa; quindi converrebbe non dar retta al navigatore e restare sull'autostrada, ma se nessuno gli dà retta l'ingorgo aumenta, eccetera eccetera.

Anche nel caso dei *prompt* si presenta una situazione simile. Il modello linguistico che permette di produrre testi è oggi addestrato sui testi scritti da umani, come Wikipedia. Ma è talmente facile produrre testi che presto le voci di Wikipedia (tra le altre cose) saranno realizzate tramite servizi di intelligenza generativa. A quel punto i testi generati automaticamente diventeranno parte dei dati sui quali verranno costruiti i prossimi modelli linguistici. Siccome però la qualità e l'attendibilità attuale dei generatori di testi derivano da quelle della base di dati usata per l'addestramento (cioè dal meccanismo di controllo incrociato degli utenti di Wikipedia), si rischierà un circolo vizioso per cui il servizio di intelligenza generativa impara dai propri errori (nel senso deleterio dell'espressione: gli errori potranno aumentare, invece di diminuire, con l'uso).

State pensando che queste discussioni però non si fanno sulle altre tecnologie usate a scuola (il mappamondo, il vocabolario, il libro)? Vero, ma si potrebbero fare. Nessuna di queste tecnologie è neutra; ognuna offre vantaggi in termini di velocità, chiarezza e condivisibilità; ognuna nasconde delle impostazioni di default che producono dei bias invisibili, di cui non ci si rende conto (come la proiezione

di Mercatore per le mappe).

Ma così si perde tempo e non si finisce il programma? Sicuro. Qui entriamo nel labirinto della libertà di insegnamento e della concezione di ogni docente del significato del termine "educazione".

Le riflessioni comuni non devono però limitarsi all'oggetto tecnologico in sé. In precedenza abbiamo indicato cinque temi, che certo non nascono con i *prompt* ma che andrebbero rilette nella luce che i servizi di intelligenza generativa gettano sopra ognuna di essi. Si tratta di temi che vanno oltre le questioni tecniche o didattiche e hanno a che fare con il modello di società che si sta costruendo, più o meno consapevolmente, e in particolare il concetto di lavoro e quello di valore.

In sintesi estrema: la maniera particolare con la quale questi sistemi eseguono il compito per il quale sono stati progettati e realizzati cambia non solo il tempo medio necessario per la loro esecuzione – e quindi il valore di questo tempo – ma anche il valore del prodotto. Questo fenomeno lo abbiamo già visto quando si sono affermati i prodotti industriali al posto di quelli artigianali: di bassa qualità, standardizzati, ma molto più economici e rapidi da produrre. Con due differenze enormi. Nelle rivoluzioni industriali del passato i nuovi prodotti erano oggetti materiali replicati a partire da uno "stampo", cioè uno schema, un progetto, un modello che veniva seguito esattamente. Lo stampo veniva riempito con materie prime, o magari lavorate, ma comunque fisiche, quindi disponibili in quantità limitata e ad un certo costo.

La rivoluzione informatica è figlia di quella ottocentesca, ma fa a meno della materia fisica (la sostituisce, potremmo dire metaforicamente, con i bit); non fa però a meno dello stampo, che è il programma. Anzi, generalizzando, potremmo dire che anche lo schema di una penna Bic è un programma che viene eseguito ogni volta che viene prodotta una delle penne, esattamente come una newsletter inviata a milioni di

destinatari è frutto dell'applicazione di uno schema. Entrambi gli schemi sono deterministici, nel senso che hanno come fine quello di ottenere copie indistinguibili del prodotto; nel secondo caso però queste copie non sono oggetti materiali ma artefatti simbolici.

Neanche i prodotti dei servizi di intelligenza generativa sono oggetti materiali; solo che per questi non è disponibile uno stampo. Non sono riproduzioni; sono produzioni, ogni volta differenti. Il risparmio qui non è calcolabile in termini di forza necessaria, ma di competenza cristallizzata. E' come se avessimo riprodotto milioni di artigiani, tutti diversi, tutti capaci di generare prodotti personalizzati a partire da una materia virtuale che è un sistema simbolico (verbale o visivo o sonoro).

Questa differenza non è casuale o secondaria, ma è proprio il fine della "terza" rivoluzione industriale, come viene chiamata quella delle intelligenze artificiali.

Dunque in un contesto educativo (che sia scuola, università, formazione professionale e non formale) si può cercare di guardare un po' avanti e immaginare cosa potrebbe succedere nei campi del lavoro, della formazione, della produzione.

1. Che succede al lavoro e ai lavori, in particolare quelli legati alla trasformazione della conoscenza? Questi strumenti si possono ancora usare come un supporto alle persone o sono stati progettati apposta per sostituirle?
2. Cosa succede alla formazione? Verrà divisa in formazione "alta", personalizzata, affidata ancora docenti umani e formazione "democratica", standard, gestita solo tramite tutor artificiali basati su grandi modelli linguistici?
3. Il costo energetico per la creazione e gestione dei servizi di generazione di opere sarà il limite che ne consentirà l'uso solo nei Paesi ricchi? Sarà riservato alla parte ricca o potente di queste società?
4. La virtualizzazione delle macchine sarà sempre più

diffusa e avrà degli effetti duraturi, ad esempio sulla piccola industria? O addirittura sulla costruzione dell'esperienza nei bambini?

5. La facilità di generazione automatica di opere a partire da prompt modificherà la maniera di valutarle in generale? Che valore (anche economico) hanno un testo o un'immagine diversi da tutti gli altri ma producibili all'infinito?

La maniera concreta di fare questo esercizio può essere differente in base al contesto (l'età, il tempo a disposizione, la capacità di interagire e far interagire).

Una lezione frontale ricca di riferimenti ed esempi?

Molte diverse ricerche che si traducono in una condivisione dei risultati?

Una discussione orale che mette in gioco i sentimenti e le opinioni?

Una scrittura personale di testi secondo modelli diversi (non necessariamente saggistici: il racconto di fantascienza si presta bene a questo tipo di esercizio)?

Una messa in scena collettiva di questi testi (a teatro, in un video, in un'animazione)?

Alcune di queste attività o tutte insieme?

Probabilmente l'idea stessa di trattare questi argomenti nel chiuso di una classe è sbagliata. Il punto di vista (l'età, la classe sociale, la lingua, la posizione geografica) sono importanti sia come punti di partenza che come punti d'arrivo. Per chi è importante il destino della formazione o il costo energetico in termini di fonti non rinnovabili? Certo non per tutti allo stesso modo.

Allora, se possibile, riuscire a costruire un dialogo mettendo insieme persone che vivono in contesti diversi potrebbe essere l'unico modo per capire l'impatto della rivoluzione del prompt per tutti.

Didattica del prompt

di Stefano Penge



Secondo le opinioni degli insegnanti intervistati dagli autori di ChatGPT [<https://openai.com/blog/teaching-with-ai>], il *prompt* si può usare a scuola in tanti modi:

1. far esercitare gli studenti in un *debate*
2. preparare quiz e lezioni
3. tradurre testi per studenti non perfettamente a loro agio nella lingua dell'insegnante
4. insegnare agli studenti come usare internet responsabilmente.

Probabilmente a noi europei, più centrati sulla nostra storia delle letteratura che sui problemi linguistici degli immigrati, questi suggerimenti sembrano ingenui e comunque non praticabili nel contesto della didattica quotidiana, a parte il secondo, che attrae irresistibilmente con la sua promessa di far risparmiare tempo al docente.

Il quarto punto (far crescere il senso critico dei ragazzi) è invece interessante. Curiosamente cita internet, e non i servizi intelligenti, come se i rischi fossero legati a

cattivi hacker in agguato là nella notte della dark web; ma è interessante proprio perché non è un suggerimento che viene da qualche agenzia culturale luddista e anti-tecnologica, ma da chi spinge per il progresso veloce e inarrestabile a qualunque costo. Sembra un'etichetta di una bevanda alcolica che recita "Bevi responsabilmente" o di un pacchetto di sigarette "Il fumo provoca il cancro".

Le prime raccomandazioni a non fidarsi troppo delle risposte dei servizi di dialogo artificiali come ChatGPT o Bard vengono infatti dai loro produttori, che sottolineano come ci possano essere "imprecisioni" nelle risposte. Dice Microsoft a proposito del suo motore di ricerca Bing potenziato con Copilot, cioè la sua versione di GPT:

"Bing mira a basare tutte le sue risposte su fonti affidabili, ma l'intelligenza artificiale può commettere errori e i contenuti di terze parti su Internet potrebbero non essere sempre accurati o affidabili. A volte Copilot travisa le informazioni trovate e potresti vedere risposte che sembrano convincenti ma sono incomplete, imprecise o inappropriate. Usa il tuo giudizio e ricontrolla i fatti prima di prendere decisioni o agire in base alle risposte di Copilot."
[<https://www.microsoft.com/it-it/bing?#faq>]

Ma qui si tratta di studenti, di ragazzi, che per ipotesi non hanno un giudizio sufficientemente solido da poter controllare i fatti (tant'è che la scuola ha per obiettivo proprio quello di insegnarglielo). Verrebbe da pensare che il limite dei 13 anni non è sufficiente per assicurare un uso consapevole di questi strumenti.

Tra i materiali a supporto della formazione Microsoft suggerisce questi obiettivi (solo in inglese, anglosassone o statunitense):

- Explain the importance of fact#checking AI-generated content.
- Describe the importance of protecting personal data and

privacy while using AI tools.

- Evaluate the potential for bias in AI-generated outputs and its implications.
 - Recognize that generative AI is a tool rather than a social companion and apply their understanding of digital wellbeing to ensure healthy digital usage
- [<https://learn.microsoft.com/it-it/training/educator-center/instructor-materials/classroom-toolkit-unlock-generative-ai-safely-responsibly>]

Gli stessi obiettivi sono declinati nelle slide in formato Power Point scaricabili dallo stesso sito. Le attività proposte mettono in campo Gina, un'appassionata di ecologia, e Alex, un giovane programmatore; viene mostrato come possano usare Bing e Microsoft Copilot ma facendo attenzione ai rischi impliciti nel loro uso.

Personalmente trovo la situazione paradossale: come se si fornissero materiali didattici sull'uso della armi, in cui però si citano i rischi legati ad un uso improprio. "Usa pure il nostro revolver, ma fai attenzione a non spararti sui piedi". E' drammatico pensare che probabilmente negli Stati Uniti esiste davvero qualcosa del genere.

D'altra parte, il cambio di direzione repentino che hanno dovuto effettuare le aziende stesse nella presentazione degli obiettivi di sviluppo delle intelligenze generative dovrebbe farci riflettere: da sostituito del motore di ricerca tradizionale (come il Multitask Unified Model promesso da Google nel 2021) sono diventate degli allegri generatori di finti acquerelli impressionisti o di riassunti di romanzi. Si può chiedere a Bing (integrato in Edge ma anche in Skype) di generare menù per una festa di compleanno dei figli, di pianificare un viaggio in Inghilterra oppure di creare una lezione sulla guerra di Secessione: insomma attività a basso rischio. L'educazione non è un contesto critico, almeno nella concezione di Microsoft.

Queste scuse anticipate vengono proposte anche in prima persona singolare, come se ci fosse un soggetto che sa di poter sbagliare, si rende conto che questo è un limite, è capace di provare vergogna, eccetera.

Ad esempio:

"I'm an AI preview, so I'm still learning. Sometimes I might say something weird. Don't get mad at me, I'm just trying to get better!"

Fa parte della retorica corrente dell'Intelligenza Artificiale: non viene presentata come un servizio (momentaneamente gratuito) gestito da un'azienda, ma come un soggetto dotato di volontà e fini.

Potremmo invece usare il prompt nella sua funzionalità più ghiotta, cioè la produzione di testi fasulli.

Se chiediamo un testo à *la D'Annunzio* che parla di una mareggiata autunnale vista da un villaggio della Maremma, poi possiamo confrontarlo con quello originale di Carducci. Così avremmo un esempio pratico della differenza tra le poetiche dei due autori ma applicate ad uno stesso tema. Magari poi potremmo fare il contrario e chiedere un testo à *la Carducci* che parla di un'improvviso scroscio di pioggia estiva durante una gita nella pineta in Versilia. In fondo è sempre Toscana.

Dicevamo che il coding è un modo di riflettere su come certe regole applicate a certi dati possano produrre o no il risultato immaginato.

Si può fare fare la stessa cosa con il "prompt" di ChatGPT?

Sì e no.

Sì, perché ovviamente si può provare a vedere cosa succede se si danno indicazioni diverse, in termini di stile o di riferimenti. Si può addirittura creare un piccolissimo dataset di testi di un certo autore, costruire un modello su quella base e vedere cosa viene prodotto.

No, perché la maniera in cui viene costruito quel modello sulla base di quel dataset e la maniera in cui viene costruito

un testo usando quel modello è del tutto opaca, si potrebbe dire sia *soggettivamente* che *oggettivamente*.

Soggettivamente: è opaca per chi la fa. Chi fa la richiesta è un utente finale, che potrà magari avere il ruolo di beta-tester ma non ha idea di cosa succeda tra il momento in cui preme il tasto "invio" e il momento in cui appare una risposta scritta sullo schermo.

Ma è anche oggettivamente opaca, cioè è opaca per tutti. La maniera di funzionare di questi software si basa sulla creazioni di ponti tra isole, ispirandosi un po' alla maniera in cui crediamo funzioni il cervello a livello cellulare. I ponti vengono creati in grande numero e continuamente disfatti e rifatti in base alla valutazione dei risultati che si ottengono. E' una strategia che non punta all'ottimo ma solo al meglio relativo, che ha senso quando il numero di isole e ponti è veramente grande, come appunto nel caso dei neuroni e delle connessioni.

E' il principio stesso del machine learning: si basa su una forma di induzione su un grandissimo numero di situazioni, catalogate con un grandissimo numero di parametri; troppi per essere elencati o controllati. E' il sogno finalmente realizzato degli scienziati inglesi del seicento: smettere di dedurre in maniera ferrea a partire da quattro idee astratte non verificabili ma finalmente costruire generalizzazioni praticamente utilizzabili a partire da dati reali. Naturalmente i vecchi empiristi lo sapevano già, che "utilizzabile" non significa "vero", e che l'applicazione del loro metodo avrebbe avuto qualche difficoltà pratica. Se si hanno a disposizione tutti i dati, allora la generalizzazione è perfetta, ma un po' inutile, perché si limita a dire quello che già si sa. Se non si hanno a disposizione tutti i dati, allora la generalizzazione è utile, ma potrebbe essere sbagliata. Questa consapevolezza è una versione applicata alla costruzione della conoscenza di altri principi regolativi che sono stati accettati da tempo in fisica: "la perfezione non è

di questo mondo". Si può avere una conoscenza imprecisa di tutto o una conoscenza dettagliata di qualcosa. Nella maggior parte dei casi, non è possibile sapere in quale delle due situazioni ci si trovi, e questa è la cosa più scoccante.

Anche qui c'è un obiettivo specificato in termini vaghi, come nell'interazione con Prolog; ma in questo caso l'obiettivo viene fornito all'interprete da qualcuno che non sa come è stato costruito il modello. Non è possibile vedere i dati (perché sono troppi) o le regole (perché non ci sono).

Che cosa si impara con queste simulazioni? Che lo stile di un autore non è niente che non si possa apprendere, a posteriori, a partire dalle sue opere. Che non serve una comprensione profonda per riprodurlo ma bastano un numero sufficiente di esempi. Che il concetto stesso di stile (e quello derivato di storia degli stili, su cui è costruita la nostra storia della letteratura) è probabilmente sopravvalutato.

Che questo sia sufficiente (che è quanto sostiene chi introduce questi strumenti in classe da "utente felice", che ha diritto di non sapere cosa c'è sotto e dietro) non è detto. E' tanto? E' poco? Probabilmente meno di quello che si imparerebbe scrivendo un programma che produce poesie usando un lessico dannunziano e secondo forme metriche dannunziane. Perché in quel caso saremmo costretti a costruire quel lessico (cosa includiamo, cosa no?) e trovare un modo per rappresentare quelle forme metriche; e poi occuparci delle metafore, delle analogie, dei colori.

Per chiudere questo paragrafo: usare ChatGPT è molto più semplice che scrivere un programma e lo possono fare tutti. Ma no, non è la stessa cosa.

Il prompt prima del prompt

di Stefano Penge



In fondo “scrivere prompt” è solo una particolare forma di programmazione. Si dice ad un computer (no, ad un software, anzi all’interfaccia di un software) quello che si vuole che faccia.

Se questo sembra lontano dalla programmazione tradizionale forse è solo perché non la si conosce a fondo e si pensa invariabilmente a una serie di stupidi comandi dati ad uno stupido robot che li esegue. Si pensa che programmare significhi specificare esattamente il risultato e il modo di raggiungerlo passo passo (l’algoritmo).

Che i prompt possano usare parole italiane, anziché astrusi comandi in un simil-inglese separati da segni di interpunzione usati a casaccio, non è nemmeno tanto strano: la storia dei linguaggi di programmazione ha già visto tanti tentativi di staccarsi dalle lingue specialistiche per avvicinare alla programmazione anche le persone normali: è il caso del COBOL, inventato negli anni ‘60 per far programmare anche i contabili e non solo gli ingegneri. L’idea poi che i linguaggi di programmazione debbano per forza assomigliare all’inglese è falsa; ci sono linguaggi basati sull’arabo, sul cinese e persino sul latino [https://www.codeshow.it/Linguaggi/Linguaggi_nazionali].

[VAI ALLA RUBRICA SULLA INTELLIGENZA ARTIFICIALE](#)

Ma nemmeno è detto che programmare sia per forza “dare ordini”. Ad esempio, nella programmazione detta “orientata agli oggetti” si definiscono le conoscenze e le abilità di un tipo di oggetti. Più oggetti insieme, collegati gerarchicamente oppure messi in relazione tramite la possibilità di scambiarsi informazioni, costituiscono un modello. Una volta costruito, il modello viene eseguito e se la vede da solo, senza interazioni con gli umani. E’ come scrivere il copione e poi mettere in scena una rappresentazione teatrale.

O ancora: nella programmazione logica non si impartiscono comandi ma si specificano tre cose: dei fatti, delle regole, un obiettivo (o meglio un teorema da dimostrare). Dati e regole costituiscono un modello. L’interprete cerca di dimostrare il teorema a partire dai fatti che gli sono noti usando le regole che ha appreso. Se ce la fa, si limita a dire “Vero”. Il terreno su cui si muove, le vie e gli incroci sono disegnati dal programmatore; ma la strada che segue non viene tracciata da nessuno.

In tutti questi casi, non vengono specificati tutti i passi da compiere nell’ordine in cui vanno compiuti, ma si fornisce una rappresentazione ad alto livello del sistema (nella programmazione ad oggetti) o degli obiettivi (nella programmazione logica). Resta il fatto che le regole secondo le quali il programma viene eseguito dal computer sono note ed eventualmente ricostruibili. Se proprio si vuole, a posteriori si può chiedere all’interprete Prolog di mostrare tutti i tentativi che ha fatto, quelli falliti e quelli che hanno avuto successo. Anche le regole “costruite” dal programma stesso durante la sua esecuzione – che è una delle caratteristiche che rende Prolog così flessibile e capace di risolvere problemi complessi – sono spiegabili secondo la grammatica che gli è stata fornita.

Software che interagiscono con le persone in linguaggio naturale sono stati immaginati e realizzati dai programmatori

già da molti decenni. A partire da **Joseph Weizenbaum**, che nel 1966 per dimostrare quanto le persone sono pronte ad attribuire coscienza a qualsiasi cosa, dalle nuvole ai calcolatori, scrive polemicamente “Eliza”, un programma che simula una sessione con uno psicoterapeuta della scuola rogersiana. Eliza non capisce nulla, ma si limita a reagire a certe parole chiave con risposte preconfezionate. Anche così, c’è chi ci passa pomeriggi interi per provare a risolvere i propri problemi.

Poi si può ricordare **Will Crowther**, un programmatore appassionato di speleologia e giocatore di Dungeons&Dragons che nel 1975 scrive “Colossal Cave Adventure”, il primo gioco testuale in cui il giocatore per esplorare una caverna infinita deve rivolgersi al suo alter-ego virtuale e *scrivere* piccole frasi come “guarda in alto, vai a destra, prendi il martello”. [https://www.codeshow.it/Codici/Colossal_Cave]

Lo stesso **Prolog**, il linguaggio principe per la programmazione logica, era stato inventato da **Alain Comerauer** negli anni ‘70 all’università di Marsiglia per scrivere programmi che permettessero di parlare al computer in francese. [<https://www.codeshow.it/Linguaggi/Prolog>]

Insomma la capacità di analizzare una richiesta scritta da una persona e produrre una risposta comprensibile è una competenza sofisticata, ma non nuova. Come non è nuova la capacità di analizzare un testo per individuarne le parti e “marcarle” in base all’argomento. Lo fanno già tutti i software dei social network quando analizzano quello che scriviamo per costruire un profilo dei nostri interessi e delle nostre opinioni, facendo quella che si chiama “sentiment analysis” (in cui “sentiment” è quell’opinione di pancia, non razionalizzata, che ci rende insopportabile la Ferragni e ci rende simpatico Gigi Proietti).

Nel caso dei modelli GPT non si sta interagendo direttamente col modello linguistico (che è un oggetto enorme con miliardi

di parametri), ma appunto con un software specializzato che è in grado di tradurre quello che scrive la persona in una serie di istruzioni che interrogano il modello. Questo software è abbastanza sofisticato da sapere dire “grazie, scusa, ho sbagliato” e individuare alcune parole importanti in un mucchio di altre.

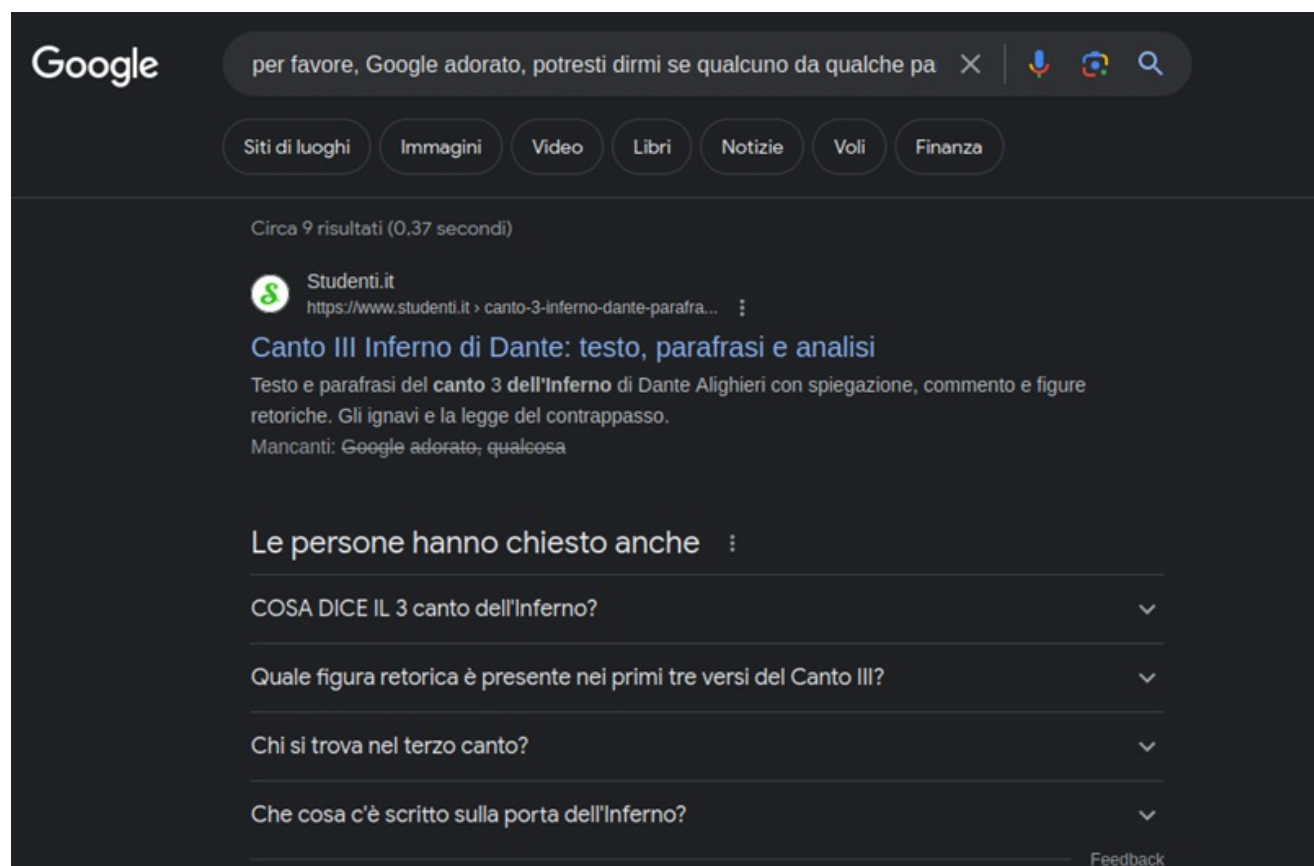
Ma non è il solo. Lo fanno anche i “prompt” dei motori di ricerca da un bel po’ di tempo; si potrebbe scrivere:

“per favore, Google adorato, potresti dirmi se qualcuno da qualche parte ha scritto qualcosa sul terzo canto dell’inferno? Grazie”

ma in pratica è sufficiente scrivere:

“terzo canto inferno”

che sono le uniche parole importanti. Il risultato è lo stesso; se non ci credete, provate.



The screenshot shows a Google search interface with a dark theme. The search bar contains the text "per favore, Google adorato, potresti dirmi se qualcuno da qualche parte ha scritto qualcosa sul terzo canto dell'inferno? Grazie". Below the search bar, there are navigation tabs for "Siti di luoghi", "Immagini", "Video", "Libri", "Notizie", "Voli", and "Finanza". The search results show "Circa 9 risultati (0,37 secondi)". The first result is from "Studenti.it" with the URL "https://www.studenti.it > canto-3-inferno-dante-parafrasi...". The title of the result is "Canto III Inferno di Dante: testo, parafrasi e analisi". The description reads: "Testo e parafrasi del canto 3 dell'Inferno di Dante Alighieri con spiegazione, commento e figure retoriche. Gli ignavi e la legge del contrappasso. Mancanti: Google adorato, qualcosa". Below the search results, there is a section titled "Le persone hanno chiesto anche" with four expandable questions: "COSA DICE IL 3 canto dell'Inferno?", "Quale figura retorica è presente nei primi tre versi del Canto III?", "Chi si trova nel terzo canto?", and "Che cosa c'è scritto sulla porta dell'Inferno?". A "Feedback" link is visible at the bottom right of the page.

In realtà l'interfaccia del motore di ricerca utilizza *tutte* le parole per eliminare possibili ambiguità; ma è in grado di filtrare il rumore di fondo costituito da articoli, congiunzioni, verbi predicativi e parole troppo comuni per concentrarsi sulle altre.

La particolarità di questa maniera di programmare computer è che lo si fa *senza specificare esattamente come*. Non è necessario specificare quello che si vuole ottenere, se non in maniera piuttosto vaga; ci pensa il software a estrarre le informazioni rilevanti e tradurle (abbiamo detto che i computer lo sanno fare bene) in istruzioni eseguibili. D'altra parte il concetto di ricerca si è trasformato parecchio: se prima era un bibliotecario a cui si forniscono le parole che devono comparire nell'oggetto cercato, oggi è un maggiordomo che propone soluzioni prima ancora che la domanda venga formulata e che il bisogno venga espresso. La trasparenza e l'anticipazione dei bisogni sono, a mio avviso, due degli obiettivi perseguiti con più costanza nel progresso dei servizi digitali. Non necessariamente a vantaggio nostro, come ho provato a raccontare qui [<https://www.stefanopenge.it/wp/trasparenze/>].

Quello che è davvero nuovo è la capacità di generare testi complessi sulla base di un modello anche stilistico oltre che di contenuto. Ed è abbastanza nuova la maniera in cui questa performance avviene: non applicando regole ad un insieme di dati, ma costruendo un enorme modello linguistico in cui le parole siano collegate ad altre parole sia per vicinanza che

per area semantica. Questo super-modello permette di predire quale parola è più probabile che segua una serie di parole precedenti. Un po' quello che fa il T-9 del nostro telefono, che ad ogni parola che scriviamo ci propone la parola successiva più probabile sulla base delle sequenze più frequenti, ma molto meglio.

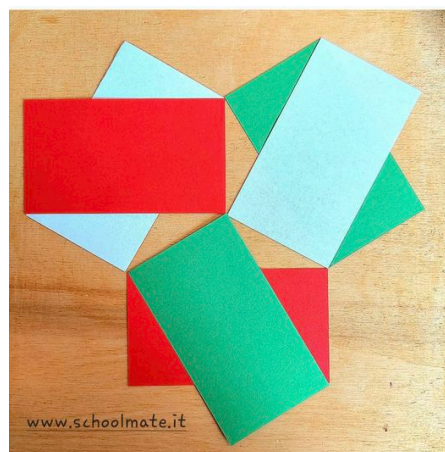
Non c'è una vera comprensione del testo, se per comprensione intendiamo qualcosa di vicino a quello che noi immaginiamo di fare quando leggiamo un testo: una specie di ricerca delle parole nel dizionario dei significati, seguita da una somma di tutti i significati trovati. Ma c'è chi dice che in fondo anche noi, quando ascoltiamo e parliamo, utilizziamo strategie superficiali che hanno a che fare più con la parentela tra le parole che con la proiezione sul piano lessicale e grammaticale dei significati di una serie di concetti. Per esempio, quando ascoltiamo qualcuno che parla una lingua nota ma non troppo tendiamo a cercare di cogliere frammenti e ricostruire il senso generale a partire da quelli e dalle intersezioni possibili tra i campi semantici relativi. Ad esempio, se intercettiamo "cane" e "grilletto" la conversazione potrebbe riguardare delle armi da fuoco; ma se insieme c'è "gatto" allora probabilmente si tratta di Pinocchio. Capito, o meglio deciso, questo contesto, tutte le altre parole verranno interpretate nella direzione selezionata. Il cane sarà Melampo, il legno quello del ciocco usato da Geppetto eccetera.

Forse addirittura i significati non sono niente altro che queste relazioni tra parole che rendono più probabili certe costellazioni di parole rispetto ad altre nelle conversazioni. Se fosse così, ChatGPT non farebbe niente di diverso da quello che facciamo noi perché è così che funziona il linguaggio in pratica.

Quindi: o noi *non* siamo più intelligenti dei grandi modelli linguistici, oppure questi modelli non sono affatto intelligenti.

Ma arriviamo all'uso didattico del prompt, che era l'obiettivo dichiarato di queste riflessioni.

Macchine che fingono di essere intelligenti



Composizione
geometrica di
Gabriella Romano

di Stefano Penge

In questi ultimi anni si sta diffondendo un approccio che epistemologicamente e eticamente sembra mettere in discussione tutto quello che abbiamo saputo o creduto di sapere. Un passaggio che vale la pena di evidenziare e di discutere, anche e soprattutto in ambito educativo.

Fino a poco fa, i computer erano al nostro fianco nell'esplorazione dell'universo, più precisi di noi, instancabili, più determinati nel non commettere errori e non trarre conclusioni avventate. Gli affidavamo il calcolo delle traiettorie degli aerei, sulla base della matematica e della

fisica che avevamo verificato nel corso due millenni precedenti, perché eravamo certi che avrebbero fatto meno pasticci di noi. Li immaginavamo come servi affidabili e incorruttibili. Un po' come il primo Terminator impersonato da Arnold Schwarzenegger: magari brutale, ma efficace e soprattutto fedele.

[VAI ALLA RUBRICA SULLA INTELLIGENZA ARTIFICIALE](#)

Invece è arrivata un'altra generazione di Terminator, più sottile, elegante, fluida. Il modello di software intelligente che va di moda oggi è bravo a fare molte cose, ma soprattutto una in particolare: *fingere*, cioè cambiare aspetto e assumere quello di chiunque altro. Non sa risolvere un problema o creare un minuetto partendo dalla regole della musica occidentale settecentesca, però può produrre degli artefatti che non sembrano artificiali, che assomigliano a qualcosa che avrebbe potuto produrre un essere umano. Di fatto, assomigliano a molte cose che esseri umani hanno prodotto, contemporaneamente.

Questa capacità, dal punto di vista della storia dei media, non è eccezionale ma è proprio tipica dei computer. E' un caso limite di *re-mediation*, cioè della tendenza dei media digitali di copiare i corrispondenti analogici, di inglobarli, di sostituirli. Ne ho parlato e scritto più volte perché è una capacità da cui sono sempre stato affascinato; due esempi qui [<https://www.stefanopenge.it/wp/strumenti-digitali-usati-analogicamente/>] e qui [<https://www.stefanopenge.it/wp/digitale-unetimologia-istruttiva/>].

L'impatto è indubitabilmente straordinario e ha fatto intravedere l'avvento della "singolarità", cioè del primo computer veramente intelligente e consapevole; tanto che mentre alcuni salutano l'alba di una nuova era transumanista in altri ambienti si sono scatenate le paure dell'intelligenza artificiale senza coscienza etica che presto dominerà il mondo

(come Skynet nella saga di Terminator).

Mentre alcuni si sono affrettati a descrivere i mutamenti nella pratica lavorativa quotidiana (almeno in certi settori: la pubblicità, la comunicazione) e a cercare di sfruttarli professionalmente in termini di consulenza e formazione, tanti docenti di buona volontà si sono impegnati a imparare a usare il "prompt", anche senza sapere bene come funziona, e si sono sforzati di capire come sfruttarlo a fini educativi; con attenzione, certo, perché altrimenti gli studenti ci mettono un attimo a farsi scrivere il riassunto del terzo canto dell'Inferno.

In generale, come per tutti i servizi digitali a cui siamo abituati, questi docenti lo usano senza fare troppo caso al fatto che il loro uso è libero e gratuito. La messa a disposizione gratuita di uno strumento che ha avuto dei costi di sviluppo e di gestione stratosferici è, a mio parere, un indicatore di un progetto di marketing più complesso, che inizia con il beta test su scala mondiale e poi passa per la commercializzazione di servizi personalizzati alle pubbliche amministrazioni, alle aziende, ai professionisti e, perché no?, alle scuole.

Parliamo allora del magico "prompt", che è il termine che si usa oggi per indicare le richieste (o ricette?) che si inviano ad un servizio online come ChatGPT per ottenere un testo (o un'immagine, o un'animazione, un audio, un video). Infatti il termine collettivo per tutti questi servizi è "intelligenza artificiale generativa".

Oggi prompt è diventato un termine di moda, una *buzz word*. E' la porta d'ingresso per l'oracolo, come direbbe Gino Roncaglia.[\[1\]](#) Tra parentesi, è curioso che dopo anni in cui la maniera di interagire con un computer usando la tastiera (anziché il mouse o il touchpad o le dita libere nello spazio virtuale) veniva considerata superata oggi è diventata una modalità avanzatissima, quasi da hacker da salotto. Nemmeno i

fieri possessori di Mac si sono lamentati. Anzi è possibile che gli effetti completi in termini di mode culturali ancora siano da vedere. Nella prossima pubblicità di Nespresso si vedrà George Clooney premere tasti per farsi fare il suo caffè preferito.

Ma che significa prompt? Come spesso succede, c'è stato uno slittamento dalla forma alla materia, dallo strumento al contenuto. E' successo per "Social Network System", che sarebbero i software che gestiscono le reti sociali ma che sono diventati "i social", cioè i gruppi di persone con cui abbiamo a che fare. E' successo per "cloud", che è il modo di dividere i software su più hardware virtualizzati, ma che è diventato il luogo unico di ogni documento che non sta sul nostro PC. Tutti questi spostamenti di senso vanno nella direzione della semplificazione, cioè della messa tra parentesi della complessità tecnica, ma anche di quella organizzativa e lavorativa. Appena i "social" divengano oggetti dotati di natura propria ci si dimentica che sono costruiti con dei filtri applicati sulla base della profilazione degli utenti, a sua volta costruita leggendo e analizzando messaggi e azioni. Appena il "cloud" diventa un sinonimo di "remoto, lontano, chissà dove" ci si dimentica che la proprietà di quel nostro documento una volta messo nella nuvola è dubbia e così sono vaghi i confini dei suoi usi possibili da parte da qualcuno che non siamo noi.

Torniamo a prompt. In inglese significa letteralmente "pronto" ed era il simbolo che appariva, e appare ancora, nelle finestre di interazione testuale con i sistemi operativi dei computer. Per esempio, in Linux può essere qualcosa di questo tipo:

```
[stefano@fedora ~]$ □
```

mentre nei vecchi computer con il sistema MSDOS appariva così:

```
C:\>_
```

In genere l'ultimo carattere lampeggia, a significare che il software è in esecuzione, non si è bloccato, ma sta aspettando un comando. E' *pronto*. A quel punto si possono scrivere comandi semplici oppure richiamare interi programmi passando dei valori come parametri. Ad esempio, in Linux il comando `ls` mostra l'elenco dei file e delle directory (come era in DOS il comando `DIR`); ma:

```
ls -d */ | grep 'Doc*'
```

fa qualcosa di più: mostra solo le cartelle il cui nome inizia per 'Doc'. Si possono anche fare anche operazioni sofisticate che agiscono su blocchi di informazioni strutturate, come questa:

```
sed -e 's/occhi/orecchi/g' -e 's/vidi/sentii/g' inferno.txt
```

che sostituisce tutte le occorrenze della parola 'occhi' con 'orecchi' e 'vidi' con 'sentii' in un file di testo che presumibilmente contiene l'inferno di Dante.

Da questo punto di vista, i prompt attuali sono un po' la stessa cosa: invocazioni di programmi (che però non conosciamo esattamente) a cui si passano come parametri una serie di valori. Ad esempio, in questo prompt:

I would like to learn about quantum mechanics. Please provide me with a brief summary of spin that could be understood by a 10-year-old.

ci sono tre parametri impliciti ("area", "argomento" e "target"); le espressioni "quantum mechanics", "spin" e "10-years-old" sono i valori che assumono questi parametri. Tutta la frase è uno schema che si può svuotare e riempire con altre triple di valori ("letteratura italiana", "donna dello schermo", "calciatore").

Da prompt è stato derivato "prompt engineering", cioè la capacità di scrivere buoni prompt, che fa sentire ingegneri

anche solo a pronunciarla. Una capacità che si insegna ormai in innumerevoli libri e corsi, alcuni dei quali dedicati a insegnanti e a studenti. Ingegnerizzare significa organizzare le cose in maniera efficiente; in questo caso, *“prompt engineering”* significa semplicemente conoscere degli schemi e saper mettere i valori giusti al posto giusto. Non sembra esattamente un'arte e non ci vuole una vita per impararla; comunque, una volta messa da parte, non si capisce bene dove possa venire riusata.

[\[i\]](#) L'architetto e l'oracolo, Laterza, 2023

Macchine che sembrano intelligenti



di Stefano Penge

Immaginate un bel seminario dal titolo accattivante: “Le macchine intelligenti e la scuola: conflitto o opportunità?”. Si presenta come obiettivo, imparziale; non è acritico ma lascia intravedere la direzione in cui occorre procedere (dalla rilevazione dei rischi alla costruzione di possibilità operative). Ottimismo della ragione e curiosità non vi

mancano, siete pronti a iscrivervi?

Un attimo. Già in questo titolo sono visibili alcuni limiti tipici dell'impostazione più diffusa.

Prima di tutto, macchine e scuola: due soggetti, due entità che inevitabilmente si confrontano. Come se ci fosse una natura artificiale che cresce, si sviluppa, produce i suoi frutti (le macchine appunto); e come se la scuola si trovasse in questo giardino incantato, novella Eva, e dovesse decidere come coglierli e mangiarli. Lieto fine: dopo qualche esitazione colpevole ogni invenzione viene assorbita dalla scuola, malgrado l'opposizione dei tecnofobi arretrati e miopi.

[VAI ALLA RUBRICA SULLA INTELLIGENZA ARTIFICIALE](#)

E' vero che non sono molte le tecnologie per l'educazione nate *dentro* la scuola, come la tavoletta di cera riscrivibile dei romani o la lavagna di ardesia (che sembra sia stata introdotta negli Stati Uniti nell'Ottocento nel quadro della formazione degli insegnanti: una tecnologia meta-educativa). Dalla tipografia alle fotocopie, dalle diapositive alla LIM: nessuna di queste tecnologie nasce per l'educazione e nemmeno per la formazione, ma tutt'al più nel quadro della comunicazione aziendale, per la diffusione rapida, economica ed efficace di informazioni.

Però l'introduzione di queste tecnologie nella scuola non è un processo naturale, continuo, unico. Si possono descrivere almeno tre modi con cui dispositivi nati in azienda sono stati successivamente usati nella scuola italiana: quello per cui sono i docenti più creativi e divergenti ad andare a caccia di tecnologie da rubare e riadattare, quello per cui è il ministero a distribuire strumenti a colpi di piani nazionali e finanziamenti su progetti e quello per cui sono le grandi

aziende alle prese con un surplus di produzione che cercano di esportare quei dispositivi nella scuola costruendo ad hoc un discorso di valori incentrato sull'innovazione, l'efficacia, le competenze del futuro. [\[i\]](#)

In poche parole: il primo passo da fare non è accettare una certa tecnologia come inevitabile per poi domandarsi come sfruttarla, ma prima occorre domandarsi da dove viene e dove si vuole che ci porti, per poi decidere se e come usarla per i fini propri dell'educazione.

In secondo luogo, penso che sia riduttivo parlare di *macchine intelligenti* contrapposte a tutto il resto. Le macchine non sono stupide o intelligenti, l'uno o l'altro; personalmente preferisco vederle come disposte in maniera continua lungo le due dimensioni dell'autonomia e della complessità. Si va da quelle semplici e pochissimo autonome – la ruota – a quelle complesse che prendono decisioni in base alle variazioni del contesto (il termostato); da quelle che possono prendere un solo tipo di decisione (agire sì/no), a quelle che hanno un set di decisioni possibili; fino a quelle che possono decidere di non agire affatto. In questo, analogico o digitale c'entra poco, né c'è bisogno di scomodare sempre HAL 9000. Vedere le ultime incarnazioni della ruota su questo sfondo aiuta a mettere a fuoco dei *temi trasversali* che a mio parere vale la pena approfondire e discutere con gli studenti: nel seguito ne indichiamo cinque.

- 1. Le macchine non sono state inventate semplicemente per fare un lavoro:** sono state inventate per fare un lavoro usando meno forza, anche se in più tempo (come la carrucola); oppure per usare una forza diversa da quella muscolare umana (come l'aratro o il motore elettrico); oppure per smontare un lavoro complesso in piccole azioni semplici che si possano ripetere e migliaia di volte, in maniera sempre uguale e senza errori (come la macchina tipografica). Negli ultimi secoli ci siamo concentrati solo sull'ultima abilità perché una volta

scoperti il motore a scoppio e quello elettrico il problema dell'uso della forza muscolare non si pone più; almeno così pare, fino al momento in cui scopriamo che le macchine che costruiscono i modelli linguistici sottostanti a ChatGPT consumano la stessa elettricità di un'intera provincia italiana. Per lo più si tratta di elettricità prodotta in loco tramite sterminati campi di pannelli fotovoltaici, quindi definita "verde"; ma gli effetti collaterali (dal consumo d'acqua nel processo di estrazione dei minerali rari come il gallio e il germanio, allo smaltimento e stoccaggio dei rifiuti) non sono banali. **Primo tema trasversale su cui appuntare l'attenzione: l'energia.**

2. **Le macchine tradizionali si possono pensare come abilità congelate in mattoncini:** abilità applicabili ripetutamente, indipendentemente da quelle delle persone che le usano. Un tipografo non deve essere uno scrittore e nemmeno un calligrafo. Staccare l'abilità dalla persona produce dei vantaggi: per esempio, le macchine possono essere usate da persone deboli o inesperte, o possono lavorare continuamente per un tempo superiore alla giornata di lavoro, senza andare in ferie o in malattia. A chi giovino questi vantaggi è un'altra questione che è legata al possesso delle macchine. La pubblicità della macchine per l'automazione dei lavori di casa presenta la casalinga come l'unica beneficiaria dell'adozione della lavatrice o dell'aspirapolvere, come le macchine per ufficio vengono presentate come aiutanti delle segretarie. Ma è ovvio che questo dislocamento delle abilità dalla persona alla macchina ha degli effetti sulle relazioni tra persone, nella famiglia come nell'ufficio o in fabbrica. **Secondo tema: i lavori che diventano obsoleti e le classi sociali che perdono potere d'acquisto.**
3. **Ha effetti – e questo dovrebbe farci riflettere per**

tempo – anche sulla formazione e sulla scuola: non ha senso insegnare alle persone delle abilità che sono di pertinenza delle macchine. Né ogni formazione professionale può essere convertita nella formazione per gestire le macchine. La macchina può finire per sostituire l'operaio, o anzi essere progettata apposta per finirla una buona volta con tutti i tipi di operai; nello stesso tempo, contribuisce alla chiusura dei corsi di formazione professionale, dalle lingue alla programmazione, e quindi permette di finirla una buona volta con tutti i tipi di docenti. **Terzo tema: il destino della formazione, per come la conosciamo ora.**

4. **Appena si sale un po' di complessità e autonomia si trovano macchine che non fanno affatto sempre la stessa cosa**, sia nel senso che eseguono una *serie* di operazioni, sia nel senso che possono eseguire alcune operazioni *al posto* di altre. Un carillon è una macchina che è in grado di eseguire sequenze di azioni (musicali) diverse, infinite sequenze di azioni. Ma può eseguire solo sequenze musicali scritte per quella determinata macchina. La stessa cosa vale per il telaio Jacquard. Entrambe sono basate sull'idea di "programma", cioè di serie di azioni scritte in maniera formalizzata, leggibili in maniera univoca ed eseguibili. Nel caso del telaio, quando si vuole avere risultato diverso, questo programma può essere sostituito. Invece di cambiare tutta la macchina (con costi e tempi enormi), si cambia solo la scheda. E' la prima volta che l'hardware si separa fisicamente dal software e i due fratelli seguono una strada diversa. Come se lo stesso corpo potesse ospitare anime diverse, a turno. Anche le macchine digitali sono in grado di eseguire azioni diverse in base al programma che viene caricato. Solo che hanno in più la possibilità di "apprendere" abilità diverse e tenerle in memoria per sempre o finché non le si mette a dormire. Un telaio Jacquard non ha memoria; invece un computer sì: quando lo si sveglia, si ricorda tutto

quello che ha appreso nel passato. Non è solo una questione di comodità: significa che un computer può eseguire un programma nel contesto di un programma precedente. In sostanza, in questo modo si reinventa il concetto di cultura – conservabile, trasmissibile, sovrapponibile – che permette di andare oltre la natura. Sembra una banalità, ma è la base per tutte le finzioni presenti e future: un computer esegue un programma in base al quale *si comporta come un altro computer*. In pratica il software si mangia l'hardware, cioè fa finta di essere un hardware diverso.

Quarto tema: la virtualizzazione.

- 5. L'enorme differenza tra un carillon e un computer è che il computer non è limitato ad un solo tipo di azioni, quindi di abilità. Ad un computer si può "insegnare ad imparare", non solo scrivendo nuovi programmi, ma inventando nuovi linguaggi con i quali scrivere altri programmi.**

Tra le cose che può imparare a fare un computer è modificare o tradurre o scrivere nuovi programmi, che in fondo sono solo sequenze di simboli con un qualsiasi testo. Si tratta per la verità di cose non particolarmente nuove, che i computer fanno da sempre e, letteralmente, in continuazione. Il fatto che sappiano trattare testi in generale, compresi quelli delle lingue naturali, non dovrebbe stupirci più di tanto. Se sapessimo che i computer passano la vita ad analizzare, riassumere, tradurre quel tipo particolare di testo che è il codice sorgente di un programma forse saremmo meno colpiti dal fatto che lo sappiano fare con l'italiano. ***Quinto tema: la produzione automatica di testi.***

Le macchine più avanzate hanno la meta-abilità di sapere quale abilità applicare e quando farlo; le potremmo chiamare

competenze. Hanno dei magazzini di risorse, e possono andarle a scegliere per usare quella più adatta. Non è escluso che la definizione standard di “competenza” come capacità di selezionare le risorse più adatte ad un problema derivi proprio da questo modo di funzionare delle macchine complesse; nel senso che quella definizione è stata scritta in un’ottica che paragona il cervello ad un computer e il pensiero al software.

Certi software, da tanto tempo, sono capaci di esercitare competenze che a noi umani sembrano “alte”, cioè vicine alle nostre: sanno giocare a scacchi, sanno risolvere problemi, in particolare quelli di ricerca all’interno di archivi sterminati. Non brillano particolarmente nel campo della produzione di oggetti nuovi ma comunque lo fanno. Se i programmatori sono in grado di definire formalmente come si risolve un problema, o come si crea un minuetto, i software sanno farlo – con piccole variazioni dovute al caso, anche quello inserito dal programmatore con parsimonia.

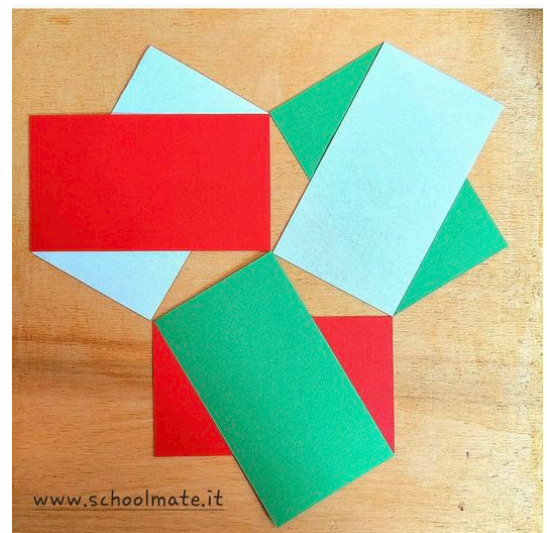
Su questa relazione programmatore-programma si basa il senso didattico di insegnare a programmare ai bambini (e in generale a persone che nella vita non faranno i programmatori, piccoli e grandi). Il famigerato e bistrattato coding è un modo di costruire dei modelli di situazione definendo le condizioni e le regole. E’ un modo di capire le situazioni, di comprendere le relazioni tra cause ed effetti, potendo sperimentare direttamente e immediatamente, correggendo le regole fino ad arrivare al risultato atteso. Non è tanto un gioco gradevole, un modo per esercitare la creatività e passare il tempo disegnando girandole e muovendo gattini, ma un laboratorio per la mente dove si mettono alla prova in maniera collettiva (e quindi discutibile) diverse interpretazioni del mondo. E’ uno dei casi a mio parere più significativi di ripensamento di una tecnologia nata altrove (la programmazione) e del suo adattamento al contesto e agli scopi della didattica. Non si fa coding per preparare futuri programmatori domani, ma per

scoprire come funziona il mondo oggi.

Questa modalità (che chiamerei “tradizionale”, a segnalare che novità e tradizione si susseguono più velocemente di quello che ci piacerebbe) viene messa in discussione dai servizi di intelligenza generativa. Il bello di questi servizi è che la competenza che esibiscono non è stata costruita a tavolino, ma è stata imparata sul campo; solo che è impossibile dire esattamente come.

[\[i\]](#)Una versione più estesa di questa descrizione è nel volume 4 di “Il mondo della rete spiegato ai ragazzi”

Perchè i docenti studiano l'intelligenza artificiale?



Composizione geometrica di Gabriella Romano

di Stefano Penge

Come in ogni occasione in cui qualche strumento viene inventato e proposto sul mercato delle professioni della conoscenza, ancora prima che parta l'ennesimo progetto nazionale per la *trasformazione delle classi tradizionali in ambienti innovativi dove gli studenti si preparino alle professioni digitali del futuro e il personale scolastico alla transizione digitale*, alcuni insegnanti si sentono in dovere di aggiornarsi e di prepararsi al nuovo che avanza. Questo va ovviamente ascritto a loro merito: invece di lamentarsi del tempo perso necessario per imparare a usare questa o quella stravaganza voluta dal ministero, invece di laudare le tecnologie naturali di una volta e rimpiangere le buone vecchie LIM, si lanciano a studiare, giocare e sperimentare coinvolgendo la famiglia, gli amici e gli studenti. Sono docenti ottimisti, entusiasti, curiosi e insieme coscienti.

VAI ALLA RUBRICA SULLA INTELLIGENZA ARTIFICIALE

Le ragioni di questo interesse didattico sono diverse: dalla difesa dei "nostri" nei mercati del lavoro internazionali alla risposta rapida alle trasformazioni culturali in arrivo, con una quota di orgoglio professionale o semplicemente un desiderio inconscio di allinearsi alla moda del momento.

Stavolta non si tratta di afferrare la meteora del Metaverso, ma di preparare gli studenti a usare l'intelligenza artificiale, cioè le interfacce dialogiche di OpenAI, cioè i prompt. I tre concetti sono spesso percepiti come uno solo: se il prompt di ChatGPT coincide con l'intelligenza artificiale, imparare a usarlo significa avere una patente che permette di entrare nel Paese dei Balocchi intelligenti.

Questo, dal mio punto di vista, non è un buon punto di partenza: ci sono mille tipi di "intelligenza artificiale", mille modi di applicarla e mille interfacce diverse. Convincersi che conoscerne *uno solo* sia la chiave per il

futuro è davvero ingenuo. E poi, anche ammettendo che si tratti della più grande invenzione degli ultimi dieci anni, niente ci assicura che la maniera specifica per averci a che fare oggi – il prompt – rimanga la stessa anche solo per i prossimi sei mesi. Avere una visione d'insieme: certo; vedere degli esempi per capire come funziona: d'accordo; imparare il modo più furbo per farsi scrivere una ricetta di cucina nello stile di Artusi... non saprei. Può essere divertente, ma quali garanzie fornisce sull'occupabilità futura (se questo era lo scopo)? Sarebbe come se negli anni '90 qualcuno avesse detto che per prepararsi alle professioni digitali bisognava imparare a usare il *mouse*. Come? Dicevano proprio così? Può darsi.

Un secondo limite che vedo in questa autoformazione è la *tonalità emotiva*. I docenti si preparano nei modi più disparati, iscrivendosi a webinar, partecipando a convegni, comprando libri, in attesa di frequentare corsi veri e propri. Accedono a flussi di informazioni di qualità diversa, proposti da attori diversi, pubblici e privati, ma che hanno in comune una caratteristica: l'adesione entusiasta e la scarsa propensione alla critica. Gli aggettivi usati per descrivere gli agenti artificiali creativi sono "entusiasmante, creativo, coinvolgente, dinamico, stimolante". Siamo ancora nella fase pubblicitaria, in cui il discente (il cliente) deve essere convinto della bontà del prodotto. Si capisce anche che nessuno abbia voglia di spendere tempo e soldi per leggere un libro che parla male della novità del momento. La modalità critica appartiene tipicamente alla terza fase, quella che viene dopo l'utilizzo quotidiano con l'apparire dei problemi. Purtroppo a volte non è possibile aspettare così tanto: a quel punto certe modalità si sono inserite nella pratica e diventano inamovibili.

Una vecchia conoscenza, un'insegnante che conosco da trent'anni, quando ho detto pubblicamente che non avevo ancora ben capito i vantaggi dell'uso di ChatGPT in classe si è

meravigliata: “ma come, proprio tu che una volta eri tra quelli più aggiornati e avanzati nel proporre tecnologie digitali?” E’ vero che quando ci frequentavamo ero uno di quelli che proponevano di *usare, smontare, fare* e non solo stare a guardare; lo facevo anch’io e mostravo i risultati dei miei esperimenti. Non è che adesso sia diventato luddista tutto insieme; ma mi pare che un’analisi critica mostri come quegli strumenti fossero più smontabili e controllabili di questi. Forse, da vero *boomer*, sono anche diventato un po’ più sospettoso sulle nuove tecnologie attuali di quanto non lo fossi sulle nuove tecnologie del passato. Ma è un atteggiamento che rivendico come conquista dovuta all’esperienza. E’ un po’ come il Lego: lo si può usare e basta, oppure ci si può domandare dove viene fabbricato e con quale plastica (se lo si fa, si scopre che i mattoncini sono fatti di Acrilonitrile Butadiene Stirene, un polimero derivato dal petrolio che ha un tempo di decomposizione stimato intorno ai 1300 anni. Noi forse non ci saremo più ma i mattoncini saranno sempre là).

Credo insomma che sia sempre utile collocare ogni nuova iniezione tecnologia nel suo contesto: da dove viene, a cosa serve, a chi porta vantaggi, cosa potrebbe diventare, quali rischi potrebbe nascondere. Più la tecnologia sembra meravigliosa, gratuita, irrinunciabile, più credo che occorra studiarla da vicino. L’Omino di burro è sempre molto abile a invitare a salire sul suo carro.

Le tecnologie non sono tutte uguali. Ci sono tecnologie che si possono smontare e tecnologie opache; tecnologie che possono potenziare l’apprendimento e tecnologie che si limitano a liofilizzarlo; tecnologie nate o almeno adattate all’educazione e tecnologie importate velocemente per massimizzare il recupero degli investimenti. Non si tratta di essere a favore o contro in generale, ma di *scegliere*.

In questa serie di articoli provo a ricostruire il senso di questa ennesima importazione nella scuola di uno strumento

tecnologico nato altrove. Lo faccio cercando di collocarlo all'interno di un contesto più generale: quello delle *macchine programmabili*.

Anche se mi sono occupato professionalmente di programmazione, di intelligenza artificiale e di didattica, non è un approccio dettato solo dalla mia storia personale e per fortuna non sono il solo a tentare di adottarlo:[\[i\]](#) anche da qualche altra parte si comincia a pensare che il rischio più grande stia proprio nel ritenere che si possa ragionare e parlare di intelligenza artificiale (in generale, di qualsiasi tecnologia della conoscenza) da *un solo punto di vista*, che sia quello epistemologico (complessità), quello tecnico (efficienza), oppure quello didattico (efficacia). Nessuna tecnologia è solo una faccenda di trovare le soluzioni migliori per risolvere problemi; per il semplice fatto che quella soluzione determinata che è incarnata da una certa tecnologia creerà problemi di altro tipo (per esempio, di sostenibilità ambientale o sociale). Così pure nessuna tecnologia può essere astratta dal contesto per cui è nata, presa in prestito e immersa nell'ambiente educativo senza considerarne le logiche profonde, senza essere ripensata e adattata.

Nella prima parte ci occupiamo delle macchine che sembrano intelligenti, poi di quelle che fingono di esserlo; vediamo poi come è nato il prompt e dove altro lo possiamo incontrare. Infine cerchiamo di capire quale potrebbe essere il senso del suo uso didattico, se ce n'è uno.

[\[i\]](#) Penso in particolare alle persone con cui ho discusso di questi temi (Rodolfo Marchisio, Marco Guastavigna, Monica Oriani)

[VAI ALLA RUBRICA SULLA INTELLIGENZA](#)

ARTIFICIALE

Intelligenza artificiale: perché pone interrogativi etici e sui nostri diritti di cittadini?



di Rodolfo Marchisio
e Stefano Penge

Dopo aver [introdotta](#) il tema, averlo [approfondito](#) vediamo di spiegare un po' di più gli **aspetti critici dal punto di vista della cittadinanza**.

Un intervento educativo (formazione di competenze e cultura digitale) è possibile solo se si *supera la visione della digitalizzazione come un processo di democratizzazione spontanea dell'accesso delle informazioni oggi non più possibile*. Agenda digitale. E come un processo "magico" e spontaneo di riforma della scuola in senso democratico grazie alle tecnologie

Quale IA

L'Europa ha stabilito i limiti da porre allo sviluppo della IA per tutelare i cittadini: [AI Act](#).

Riguardano i *modelli fondativi alla base di grandi sistemi di AI* e il ricorso alla sorveglianza biometrica e alla (ipotizzata) *polizia predittiva*. L'AI Act, **inquadra i diversi sistemi di intelligenza artificiale** pone paletti, **proibisce alcune applicazioni** e introduce procedure di salvaguardia per

mettere al riparo i cittadini dell'Unione da **abusi e violazioni** dei diritti fondamentali. L'attenzione è:

1- sui **modelli fondativi**, quelle forme di intelligenza artificiale generali in grado di svolgere compiti diversi (come creare un testo o un'immagine) e allenati attraverso un'enorme mole di dati non [categorizzati](#).

Si è lavorato ad una applicazione preventiva delle regole su [sicurezza](#) informatica, **trasparenza dei processi di addestramento e condivisione della documentazione tecnica** prima di arrivare sul mercato. Wired.

2- La UE è arrivata ad un [compromesso](#) sull'uso dell'AI per **compiti di polizia e di sorveglianza**.

Sul riconoscimento biometrico in tempo reale, ma si era discusso anche di [polizia predittiva](#) poi vietata. Ossia usare gli algoritmi per prevedere le probabilità con cui **può essere commesso un reato, da chi e dove?** Quali diritti verrebbero violati con questa delega? [AI Act Europa](#) (Wired)

Oltre a questo ci sono nodi e modi di vedere la IA (che è mondo complesso) che emergono anche dal fatto che il termine sia sempre più spesso accompagnato da un aggettivo (generativa, etica, spiegabile ...).

Si riflette su questi **temi e interrogativi**:

- L'intelligenza artificiale può **prendere decisioni** basate sui **dati** e sul **passato**. Talora **errate** a causa di una mancanza di consapevolezza del contesto. Esiste una possibilità di sapere come? Dovrebbe essere la [IA explainable](#). (spiegabile). Ma funziona? Per tutta la IA?
- Una decisione può influenzare la vita delle persone, essere utilizzata per scopi maligni, essere **influenzata dai pregiudizi e dai valori dei suoi creatori**. Gli algoritmi utilizzati nell'IA possono essere influenzati da **pregiudizi, anche involontari**. O "premiati" in modo da imparare in una certa direzione prefissata da chi li progetta. Ciò può portare a decisioni discriminatorie, sbagliate, diverse da quelle che avrebbero preso degli

umani (processi di selezione del personale, prestito bancario...). Anche se le stesse decisioni umane sono **diverse se prese da persone competenti, da potenti, da una maggioranza più o meno informata**. Come constatiamo da anni. Ma il meccanismo si potenzia ed opacizza con IA.

- L' IA comporta la raccolta e l'elaborazione **di grandi quantità di dati personali**. Questo solleva preoccupazioni sulla **privacy** delle persone e sulla **sicurezza dei dati**.
- L'IA sta automatizzando e **sostituendo sempre più lavori umani**. Ciò solleva preoccupazioni sulla perdita di posti di lavoro e sulla necessità di una riqualificazione delle persone.

Vedi [studi del possibile impatto sulla occupazione](#).

- Nonostante i progressi nell'IA, esistono ancora criticità sui sistemi **di controllo decisionale affidabili e trasparenti**. Questi problemi richiedono una **rigorosa regolamentazione, un'etica e una responsabilità riguardo all'uso dell'IA, nonché un'attenzione costante alla sua implementazione e sviluppo**.

Stefano Penge, filosofo e informatico, avverte:

“L'interesse recente – da un anno a questa parte – per l'intelligenza artificiale, con tutti i discorsi a favore e contro, si è concentrato su un sottoinsieme particolare di abilità che attraggono e spaventano apparentemente più delle altre: quelle linguistiche, che permettono ad un software di analizzare un testo, di tradurlo, di prepararne un riassunto, ma anche di continuarlo o riscriverlo secondo un altro stile; il tutto, attraverso un'interazione continua con le persone (il famoso 'prompt').

*Il riconoscimento facciale e la guida autonoma (le altre abilità di cui si occupa l'AI ACT) hanno in comune con questo sottoinsieme linguistico l'abbandono della **logica tradizionale**, quella basata su regole e deduzione, a favore di una **logica induttiva**. Niente di nuovo, insomma: tutto sommato*

anche la nostra specie funziona così la maggior parte del tempo e solo in casi speciali usiamo la logica classica e il ragionamento formale. Vedi anche <https://www.stefanopenge.it/wp/intellig-enti/>.

Quelli citati sopra sono tutti esempi di **machine learning**, che indica una maniera di costruire modelli artificiali che simulano una parte del mondo (per esempio, un testo) a partire dalla raccolta di dati relativi a miliardi di situazioni, per induzione appunto.

Questo sistema è **inerentemente soggetto ad errore**, perché i dati di partenza non sono tutti i dati ma solo una selezione; perché la maniera in cui sono stati selezionati quei dati (e la maniera in cui vengono “premiati” i modelli migliori) potrebbe essere non oggettiva ma guidata da preferenze culturali o sociali. Ma soprattutto perché il futuro non è per forza uguale al passato, mentre il machine learning per sua natura si basa proprio su questa fede.

Ma anche se decidessimo di fidarsi delle competenze dei software basati sul machine learning, il problema nasce quando queste competenze vengono usate non a fianco, ma **al posto** di quelle umane come se fossero equivalenti in qualità, **perché costano di meno e rendono di più a chi le governa.**”

Anche **Sasha Luccioni**, ricercatrice nell'AI etica e sostenibile -HuggingFace – sottolinea:

1. **Consumi energetici troppo elevati, i bias cognitivi che vengono riproposti dagli algoritmi, e la tutela del copyright sono le tre sfide che [l'intelligenza artificiale dovrà affrontare](#).**

La diversità deve essere presente all'interno dei modelli dell'intelligenza artificiale: “La tecnologia non è neutra, i bias dell'[intelligenza artificiale](#) altro non sono che quello che noi vediamo nella società. **Qualsiasi stereotipo si applichi viene peggiorato dal sistema, che tende ad amplificare la distorsione. Solo il 12% di chi lavora nell'AI è donna** ...è fondamentale che

questa tecnologia *“non sia nelle mani di pochi soggetti, deve essere open, per intervenire su questi aspetti in modo condiviso”*. Wired

2. **Per tutelare la proprietà intellettuale e il copyright** (Luccioni) bisogna *“implementare meccanismi per capire quali siano le fonti che sono state utilizzate, analizzare i **dataset** usati da un’intelligenza artificiale*. Un lavoro immane che va fatto fare dalle macchine, ma **come e da chi controllate?** E i cittadini/utenti in tutto questo?
3. **L’impatto ambientale non è trascurabile:**
Ad esempio, sono state emesse 500 tonnellate di CO₂ per la creazione di Gpt-3”. Wired

A proposito della possibilità di controllare l’IA Penge paragonava la IA ad una piccola bomba atomica: ha un potere enorme, ma può essere costruita (e smontata) solo da quelli che hanno sufficiente potere per raccogliere, selezionare e gestire i materiali (radioattivi!) necessari. Il semplice cittadino non avrà mai questa possibilità; al massimo può giocare con le interfacce che permettono di interagire con i software di IA.

“Per dire meglio ci sono tre aspetti:

- uno è quello della **chiusura del software**, in questo caso della **proprietà** dei modelli linguistici utilizzati, che hanno grande valore perché sono stati prodotti con molte ore di calcolo e molto lavoro di revisione umana;*
- il secondo è quello delle **dimensioni di questi modelli**, che per essere creati e trattati richiedono risorse informatiche enormi, che consumano elettricità, scaldano, etc.;*
- il terzo è quello dell’**opacità**: un Large Language Model è costruito con miliardi di parametri. Significa che **non è analizzabile da umani, non si può capire perché risponde in un modo o in un altro**. Questi tre aspetti insieme **impediscono di mettere le mani in questi***

***strumenti** come si potrebbe fare con un software libero, ad esempio con LibreOffice (ma non con MS Word)."*

Tornando alla scuola un esempio di preoccupante ottimismo in questo [podcast](#) che ipotizza una scuola alla cinese. Paese che notoriamente è da tempo l'incarnazione del "grande fratello" che viola costantemente i diritti dei cittadini, sudditi [ipercontrollati](#) e prigionieri. Concluderemo sulla scuola.