

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 letterature 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.