



Come comunicare le AI generative senza generare false credenze

di Giovanni Acerboni, 28 giugno 2024

Nicholas [Belmore](#), ricercatore di Harvard, mi ha chiesto, in una conversazione su [Linkedin](#), che cosa ne penso sulla possibilità di trovare per le AI generative un linguaggio affidabile e nello stesso tempo comprensibile ai non esperti (“both reliable and accessible”). Secondo Belmore, infatti, molte espressioni utilizzate dai data scientists (informatici, matematici, ingegneri ecc.) creano confusione nel grande pubblico: intelligenza (artificiale), ragionamento, allucinazioni ecc.

Gli ho risposto che in un post mi sarebbe stato difficile rispondere adeguatamente e mi sono impegnato a presentare le mie idee in un documento più strutturato. Questo.

1. Le nuove parole

Chi inventa o scopre o crea qualcosa, la nomina creando un neologismo. I neologismi si creano in diversi modi, ma quello che interessa qui è il **neologismo semantico** (o risemantizzazione o rideterminazione), che consiste nell'**attribuire un significato nuovo a una parola esistente**.

Molto spesso il nuovo significato ha una qualche relazione semantica con la parola esistente. Alcuni esempi dagli ambiti digital, hardware e software: web, (web)page, home(page), (web)site, mouse, application, program.

Il significato di questi neologismi può creare disorientamento, soprattutto nei primi tempi della loro comparsa. Infatti, quando il lettore incontra per la prima volta, per esempio, '(web)page' vede subito che ci sono alcune analogie ma anche molte differenze con il significato a cui è abituato. Per esempio, le pagine web non sono numerate, non appaiono una dopo l'altra in una sequenza vincolata, possono essere modificate e persino cancellate.

Quando i neologismi riguardano COSE, virtuali (page) o tangibili (mouse), i fraintendimenti durano poco, i significati nuovi si stabilizzano e nessuno fa più confusione.

La situazione è molto diversa **quando dalle COSE si passa a CONCETTI ASTRATTI** come, per stare nel nostro ambito, intelligenza, ragionamento, creatività, conoscenza, coscienza, argomentazione, persuasione, manipolazione ecc.

Sul significato di questi concetti l'umanità si interroga da un paio di migliaia di anni: filosofi, teologi, logici, biologi, neurologi, chimici, cognitivisti, linguisti, ecc.

Un significato unico e condiviso di termini come questi è raro e in ogni caso le ricerche sono sempre in corso e spostano sempre in avanti il limite della conoscenza. Sicché, **termini come questi hanno significati instabili** che nell'ambito di uno specifico discorso specialistico vengono ben definiti e dunque risultano chiari agli esperti.

I non esperti invece utilizzano questi termini alla buona, vi attribuiscono un significato generico, valido per le conversazioni non tecniche dove gli errori terminologici non contano molto e le incomprensioni si superano facilmente.

Ma ogni volta che un non esperto incontra uno di questi termini in un discorso specialistico non sa soppesare quanto il significato a lui noto differisca dal significato attribuito dallo specialista, a meno che lo specialista non glielo spieghi molto chiaramente (cosa non facile, che accade infatti raramente).

Nella comunicazione con non esperti **il significato di questi termini è dunque molto spesso indeterminato e di conseguenza è ambiguo**, cioè può essere interpretato in modi diversi.

L'effetto di una comunicazione ambigua è sempre negativo. Il destinatario può:

- non capire;
- capire un'altra cosa;
- eseguire un comportamento scorretto;
- diffondere un'informazione scorretta;
- dubitare delle proprie competenze e conoscenze;
- dubitare dell'affidibilità chi gli sta parlando, fino a considerarlo un ipocrita.

2. Un linguaggio per le AI generative

C'è un peccato originale: l'Intelligenza Artificiale si autodefinisce nel 1956 con un termine (Intelligenza) dal significato indeterminato. L'aggiunta di Artificiale restringe l'ambito del discorso ma non leva l'indeterminatezza.

Ciò è accaduto verosimilmente a causa dell'ipotesi che una macchina potesse pensare e agire come un uomo. Buona parte della ricerca dell'AI tenta di riprodurre il meccanismo del ragionamento e del comportamento umano. Per esempio, il termine Reti Neurali deriva da questo approccio.

Ad ogni modo, è **insensato contrastare i termini che si sono ormai stabilizzati**, come appunto Intelligenza Artificiale. Non si può non usare questa espressione.

Diverso il discorso per i termini che non designano concetti tecnici stabilizzati, ma ne derivano per l'appartenenza alla stessa area funzionale e semantica (nei significati tradizionali), come ragionamento, creatività ecc. che infatti vengono spesso utilizzati tra virgolette (chi le usa sa e vuole far sapere che le usa in modo improprio, come scorciatoia per intendersi facilmente con gli interlocutori).

Di tutti i termini non tecnici si può evitare la stabilizzazione: basta sostituirli con altri migliori.

Tuttavia, fino a quando il discorso rimane confinato agli esperti, il tema non si porrebbe nemmeno. **Il tema invece si pone e in modo estrapolatamente critico quando il discorso raggiunge i non esperti.**

È il caso delle AI generative, soprattutto quelle del linguaggio, perché hanno tolto all'uomo l'esclusiva dell'uso efficace del linguaggio. Il che è sorprendente, soprattutto per i non esperti di AI.

I non esperti di AI sanno tuttavia che cosa sia un linguaggio, la comunicazione, la scrittura. Sanno che **il linguaggio esprime sempre un contenuto** (nonostante Chomski la pensi diversamente), e la scelta dello stile connota funzionalmente quel contenuto. Il linguaggio è dunque il prodotto comunicativo di un ragionamento, cioè esprime l'intenzione di comunicare qualcosa a qualcuno.

Utilizzando un software che scrive i non esperti si domandano come diavolo faccia. Se gli si risponde con i termini che valgono per la comunicazione tra umani, si fa un gran pasticcio. Perché un software è un software, non ha intenzioni, dunque non può condurre un ragionamento su quale sia il contenuto da comunicare e come. Per scrivere, **segue altre logiche. Per spiegarle, le metafore sono delle vere trappole.**

3. Norme tecniche

Le regole per la comunicazione precisa e comprensibile di un discorso tecnico esistono. Long story short:

- sintesi: comunicare solo le informazioni necessarie, cioè utili a capire e utilizzabili per eseguire un comportamento;
- chiarezza: costruire frasi brevi e utilizzare termini comuni e nello stesso precisi. Dei tecnicismi che non possono essere evitati va fornita una traduzione in termini comuni.

Queste regole sono state formalizzate nella legislazione tecnica.

In Italia abbiamo dal 2013 lo Standard 11482:2013 [Elementi strutturali e aspetti linguistici delle comunicazioni scritte delle organizzazioni](#) emanato dall'UNI, che è la filiale nazionale dell'ISO.

L'ISO ha emanato nel 2023 lo Standard 24495-1 [Plain language. Part 1: governing principles and guidelines](#) e sta elaborando la [Part 2: Legal communication](#) e la [Part 3: Science writing](#).

4. Ragion Informatica Ben Temperata

Non è sempre facile per uno specialista ottenere uno stile efficace per la comunicazione con i non esperti. Gli specialisti infatti:

- temono di dire una cosa diversa se cambiano stile (preoccupazione legittima)
- credono, a torto, che i loro contenuti non possono essere comunicati diversamente
- faticano ad abituarsi a uno stile diverso

Dipende sempre dallo scopo e dalle motivazioni (personali o aziendali). Chi vuole, può. Tanto più che le Intelligenze Artificiali supportano la semplificazione. **Il Natural Language Processing funziona ottimamente**, perché automatizza l'identificazione delle frasi e delle parole da semplificare. Per la lingua italiana ne avevo realizzato uno io (in collaborazione con [Alessandro Panunzi](#) dell'Università di Firenze), con algoritmi brevettati utilizzati nel machine learning e un database con migliaia di termini. Ma attualmente non è in funzione. Ce ne sono altri per altre lingue.

Quanto alle AI generative, non sono molto brave a riconoscere i fatti linguistici complessi, però è innegabile che scrivono in modo piuttosto chiaro, salvi gli errori di contenuto, che sono piuttosto frequenti e gravi quanto più il discorso è specialistico.

Ho scritto "errore". Gli errori sono errori, non sbagli e tanto meno allucinazioni. Se c'è un errore a valle, c'è un errore a monte. Potrà essere controintuitivo, ma **la riparazione degli errori avviene riducendo il dataset, non aumentandolo**.

Sarebbe utile un corso di filologia e di pragmatica nelle lauree STEM e nelle aziende che fanno machine learning.