

CLINFO.AI:

RICERCA BIBLIOGRAFICA E CONFRONTO CON PUBMED.

L'attività di *reference* in biblioteca riferita alla richiesta di supporto nella ricerca bibliografica da parte degli utenti di medicina (studenti, docenti, ricercatori e clinici) è destinata a subire un cambiamento radicale in tempi molto brevi, grazie alla diffusione dei modelli linguistici di grandi dimensioni (**LLM**) (large Language Mode) come ChatGPT introdotta nel 2022.

Infatti, l'applicazione di questi algoritmi è entrata in modo dirompente anche in un ambito che, fino a pochissimo tempo fa, era di dominio solo umano. I chat bot ora sono in grado di selezionare e identificare in modo efficiente gli studi che soddisfano criteri specifici e semplificare la revisione della letteratura fornendo anche dei *summary* originali.

PubMed è un data base ad accesso libero, prodotto dalla National Library of Medicine (NLM), che offre la possibilità di ricercare nell'archivio MEDLINE le citazioni bibliografiche di ambito biomedico. Nel 2024 la NLM ha introdotto un nuovo algoritmo per la tempestiva indicizzazione dei termini MeSH delle citazioni in MEDLINE: il nuovo algoritmo MTIX (Medical Text Indexer-NeXt Generation).

Si tratta di un modello di apprendimento automatico noto come rete neurale, un tipo di intelligenza artificiale.

Noi bibliotecari di area medica eravamo chiamati a condurre ricerche bibliografiche all'interno di data base specialistici, primo fra tutti PubMed, tenendo in considerazione le caratteristiche del data base, le modalità di indicizzazione del Medline e ovviamente i termini MeSH. Affiancavamo gli utenti nel costruire *query* di ricerca per identificare e selezionare la letteratura pertinente ad un quesito. Il professionista poi avrebbe analizzato e selezionato la letteratura di suo interesse.

Dalla fine del 2022 è arrivata l'intelligenza artificiale generativa. ChatGPT e simili hanno stravolto il "mondo" in un modo mai avvenuto prima, basti pensare che gli utilizzatori di ChatGpt-3.5. sono stati di un milione di utenti dopo soli 5 giorni dal lancio avvenuto il 30 novembre del 2022. (1)

ChatGPT (acronimo di Chat Generative Pre-trained Transformer. "Trasformatore generativo pre-addestrato") è una chatbot basata su intelligenza artificiale e apprendimento automatico, sviluppato da OpenAI un laboratorio di ricerca sull'intelligenza artificiale costituito dalla società no-profit OpenAI e dalla sua sussidiaria for-profit OpenAI fondata nel 2015 a San Francisco, tra i fondatori c'è Elon Musk. È pensato per interagire nella conversazione con un utente umano. (2)

Si tratta quindi di applicativi capaci di creare un rapporto di fiducia uomo - algoritmo.

A conferma di questa tendenza i Big Tech stanno scommettendo miliardi di dollari proprio sul fatto che noi umani ci fidiamo abbastanza della sicurezza di questi sistemi e anche sulla loro capacità di gestire anche le nostre informazioni private.

Immaginiamo di utilizzare l'archivio continuamente aggiornato di Medline e indicizzato con il nuovo algoritmo MTIX, aggiungiamo la piattaforma Github open source che offre un servizio di hosting per progetti software e utilizziamo l'algoritmo di intelligenza artificiale addestrandolo su un set di dati (PubMed Retrieval and Synthesis (PubMedRS-200) pubblicamente disponibile di 200 domande strutturate in formato Open QA, abbinate a risposte derivate da revisioni sistematiche (che seguono le linee guida del PRISMA)(3) e riferimenti corrispondenti, otteniamo Clininfo.ai, una WebApp open-source molto promettente per uno sviluppo futuro.

Il sito si presenta nel seguente modo:

“Welcome to the official repository for **Clinfo.ai**: An Open-Source Retrieval-Augmented Large Language Model System for Answering Medical Questions using Scientific Literature”

Clinfo.AI ricerca e sintetizza la letteratura medica su misura per uno specifico quesito clinico per fornire una risposta basata sulla letteratura indicizzata. Sfruttando una catena di clinfo.ai LLMS, è in grado di analizzare il contesto dell'indagine per identificare e presentare gli articoli più rilevanti pertinenti a una domanda scientifica” Inoltre consente di personalizzarlo aggiungendo set di dati.

Al seguente link è possibile vedere la demo:

[GitHub - som-shahlab/Clinfo.AI: This is Clinfo.AI Demo Instruction](https://github.com/som-shahlab/Clinfo.AI)

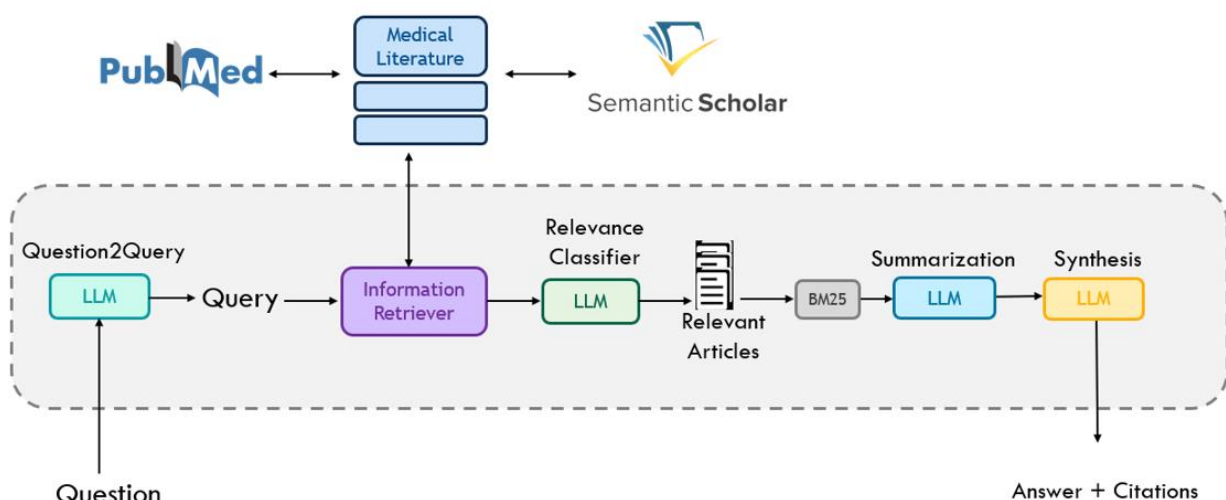


Fig.1

Come evidenziato nella figura 1, il *chat bot* può lanciare la *query* anche in Semantic Scholar. (4)(5)

Nel sito della web app si evidenziano bene le differenze fra una domanda in Clinfo.ai e chat GPT4, ma a noi bibliotecari interessa maggiormente il paragone con una *query* equivalente in PubMed.

Nell'articolo di Alejandro Lozano ((6) riportato in bibliografia vengono mostrati dettagliatamente il funzionamento delle *query*, le risposte dell'algoritmo e l'accuratezza della sintesi. Per approfondire di seguito il link alla pubblicazione:

[Clinfo.ai: An Open-Source Retrieval-Augmented Large Language Model System for Answering Medical Questions using Scientific Literature \(worldscientific.com\)](https://www.worldscientific.com/doi/10.1142/S0191261523501401)

Tralasciando gli aspetti strettamente informatici, la novità di Clinfo.ai più interessante sta nel fatto che la macchina è in grado di creare autonomamente una *query* con i termini mesh a patto però di impostare una domanda non generica ma in stile PICO. (7)

Inoltre, se facciamo alcune prove nella web app, ci accorgiamo che se scegliamo il data set di PubMed, piuttosto che il data set di Semantic Scholar, non solo la *query* terrà conto dei termini MeSH ma anche la selezione della letteratura sarà ottenuta con un taglio di risposta per professionisti, mentre se lanciamo la *query* in Semantic Scholar l'indicizzazione nella scelta delle citazioni bibliografiche e il riassunto terranno in considerazione che il possibile interlocutore umano non è un professionista.

Nell'esempio che propongo ho impostato una *query* in PubMed e in Clinfo.ai e il risultato è stato sorprendente.

La stringa di ricerca è stata pensata per un argomento di "nicchia" e con un numero di citazioni limitato per verificare agevolmente e per effettuare il confronto e la successione delle ricerche è impostata partendo prima da una ricerca con parole chiave e poi con i termini MeSH.

Materiali e Metodi

Step:1

Ricerca i termini MeSH e restringere la ricerca con MeSH Major Topic.

- Deglutition disorders; Caregivers
- Medication Therapy Management

Step:2

Lanciare le *query* in PubMed:

- Prima Query

Parole chiavi con operatori booleani, non sono stati inseriti filtri:
dysphagia AND caregivers AND management

Risultato: 187 citazioni

- Seconda Query

Termini mesh e operatori booleani, non sono stati inseriti filtri:
("Deglutition Disorders"[Majr]) AND "Caregivers"[Majr]

Risultato: 42 citazioni

- Terza query **Generata da Clinfo.ai:**

"Deglutition Disorders"[MeSH Terms] OR "Deglutition Disorders"[All Fields] OR
"swallowing disorders"[All Fields] AND ("Caregivers"[MeSH Terms] OR
"Caregivers"[All Fields])

Risultato: 290 citazioni

Aggiunta dei filtri per tipologia di documenti: Clinal trial; Meta Analysis e Systematic review

Risultato: 21 citazioni

Allegato A: elenco delle 42 citazioni in PubMed

Step: 3

Lanciare la ricerca in Clinfo.ai

- Primary Queries Used:
("Deglutition Disorders"[Majr]) AND "Caregivers"[Majr]

Risultati: 15 articoli

- Queries Used **Generata:**

("Deglutition Disorders"[Majr]) AND "Caregivers"[Majr]

("Deglutition Disorders"[Mesh]) AND "Caregivers"[Mesh]

("Deglutition Disorders"[Mesh] OR swallowing disorders) AND

("Caregivers"[Mesh] OR caregivers)

Clinfo.ai genera in autonomia una query più ampia, senza il limite del Major topic, inoltre utilizza il termine *swallowing disorders* che non è un termine MeSH ma una parola chiave molto usata per indicare i problemi di deglutizione.

Se ri-lanciamo la ricerca nello stesso modo, la seconda volta i risultati sono ancora più interessanti perché la web app si ricorda della prima query e aggiunge una stringa ulteriore di ricerca, non è soddisfatto dei risultati che ha ottenuto, troppo limitati all'oncologia pediatrica?! Non sappiamo cosa “pensa” però aggiunge una *Query* lanciandola su tutti i campi [All Fields] invece che limitarla con [MeSH Terms]

Esattamente la stessa cosa che avrebbe fatto un umano: provare differenti *query* fino a identificare quella che risponde maggiormente al quesito di *scanning* o *searching* e focalizzare meglio la domanda di *background* o di *foreground*. (8)

I risultati delle ricerche sono riportati negli allegati:

Allegato B: Summary e citazioni in Clinfo.ai

Allegato C: Summary e citazioni selezionando Semanti Scholar.

- **Second queries Used:**

"Deglutition Disorders"[Majr]) AND "Caregivers"[Majr]
 ("Deglutition Disorders"[Majr] OR "swallowing disorders") AND
 ("Caregivers"[Majr] OR "caregiving")
 ("Deglutition Disorders"[MeSH Terms] OR "Deglutition Disorders"[All Fields] OR
 "swallowing disorders"[All Fields]) AND ("Caregivers"[MeSH Terms] OR
 "Caregivers"[All Fields])

Step 4

Verificare gli articoli intercettati da entrambe le ricerche e confrontarli con i risultati di Clinfo.ai. La ricerca è stata fatta su tutti i data base in data 30 luglio 2024.

| DATA BASE | | | | | | | | |
|----------------------------------|--|---|---|---------------------------------------|--------------------------------|-------------------------|---|--|
| | PUBMED | CLINFO.AI | CLINFO.AI | PUBMED + Filtro Revisione sistematica | COCHRANE | SEMANTIC SCHOLAR | SEMANTIC SCHOLAR | PUBMED |
| | ALLEGATO A | ALLEGATO B | QUERY AUTOGENERATA | | | ALLEGATO C | | |
| Ordinamento citazioni: Best Mach | PUB MED: ("Deglutition Disorders"[Majr]) AND "Caregivers"[Majr] + 42 citazioni | CLINFO.AI ("Deglutition Disorders"[Majr]) AND "Caregivers"[Majr] + 15 citazioni | QUERY: "Deglutition Disorders"[Majr]) AND "Caregivers"[Majr] ("Deglutition Disorders"[Majr] OR "swallowing disorders") AND ("Caregivers"[Majr] OR "caregiving") ("Deglutition Disorders"[MeSH Terms] OR "Deglutition Disorders"[All Fields] OR "swallowing disorders"[All Fields]) AND ("Caregivers"[MeSH Terms] OR "Caregivers"[All Fields]) | PUBMED- Filtro Revisioni | COCHRANE TRIALS = 10 citazioni | Semantic SCHOLAR: query | Semantic scholar- Randomized Controlled Trial | Query generata da Clinfo in PUB MED con filtri (Systematic Review;RCT;Meta Analysis; Clinical.Trial) +21 citazioni |
| PMID | | | | | | | | |

Fig.2 Data Base e query relative.

Risultati:

I 15 articoli selezionati e indicizzati da Clinfo.ai sono gli stessi che risultano in PubMed indicizzati come i primi 15 nell'elenco delle citazioni, la corrispondenza è del 100%. Inoltre, possiamo vedere come Clinfo.ai è stata in grado di includere nell'elenco delle citazioni le Systematic Review e i Trial, presenti nella Cochrane data base.

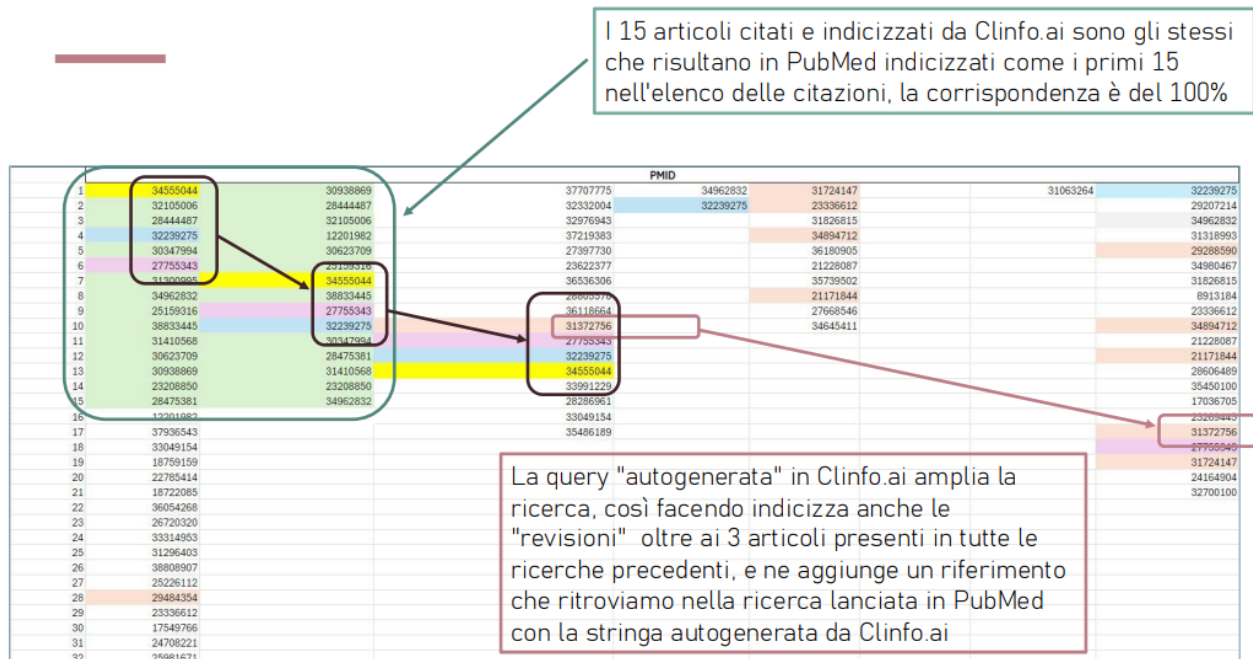


Fig.3 In verde le citazioni che coincidono in PubMed e Clinfo.ai

Legenda dei colori:

| | |
|--|--|
| | In verde le citazioni che si sovrappongono nei due data base |
| | In giallo le citazioni che si sovrappongono con la query autogenerata |
| | In violetto le citazioni che si sovrappongono tipo di documento Revisione Sistemática |
| | In azzurro le citazioni che si sovrappongono nelle diverse query anche in PUBMED anche con filtro attivo |
| | In arancione le citazioni presenti in anche nella Cocrahe |
| | In rosso le citazioni che si sovrappongono fra Pubmed e semantic Scholar |

Conclusioni

La letteratura selezionata e indicizzata da Clinfo.ai è assolutamente pertinente, inoltre, gli articoli che seleziona per fare la *summary* e che riporta in bibliografia coincidono con quelli indicizzati in PubMed.

In conclusione, come ben evidenziato nell'articolo di Lonzano “vale la pena notare che studi precedenti hanno riportato che gli LLM dimostrano la capacità di generare operatori booleani e sintassi accurati, aderendo efficacemente ai formati di query PubMed. Tuttavia, nell'articolo si sottolinea che questi modelli generavano anche termini MeSH “allucinati”, che potrebbero potenzialmente portare all'esclusione di studi pertinenti”. (6)

Aggiungerei però anche che gli errori generati dal LLM sono misurabili e quindi come tali di possibile soluzione ecco perché è mia opinione aspettarsi una realise molto performante in tempi brevissimi.

Link ai materiali supplementari:

Confronto Citazioni Clinfo.ai e PubMed:

https://liveunibo-my.sharepoint.com/:x:/g/personal/simonetta_righi_unibo_it/ESxIgh0vTIRJnFjoYDFfWKABHvgiD71lzuYxW9G8lRA37Q?e=YDsE1k

Allegato A - Elenco delle 42 citazioni in PubMed

https://liveunibo-my.sharepoint.com/:x:/g/personal/simonetta_righi_unibo_it/EZxaTAV1nhNBI4hUN3SENG8BEIRcasSQFALHtzRAqHTOpQ?e=fAA8c7

Allegato B - Summary e citazioni in Clinfo.ai

https://liveunibo-my.sharepoint.com/:w:/g/personal/simonetta_righi_unibo_it/EfYfUx7jt15OifkN60SKPbMBtEUkbfzNT0dGYxVBXdUkzg?e=crbvbf

Allegato C - Summary e citazioni selezionando Semantic Scholar

https://liveunibo-my.sharepoint.com/:w:/g/personal/simonetta_righi_unibo_it/EXgZOSxuazBNj5FHPBHm5kABHItfY7JNc-1CIEJro24qDQ?e=KZFufy

BIBLIOGRAFIA

1. Il primo anno di ChatGpt: la storia della chatbot che ha inaugurato l'era dell'intelligenza artificiale generativa. Corriere della sera.
2. Introducing ChatGPT | OpenAI [Internet]. [cited 2024 Aug 7]. Available from: <https://openai.com/index/chatgpt/>
3. PRISMA 2020 — PRISMA statement [Internet]. [cited 2024 Aug 7]. Available from: <https://www.prisma-statement.org/prisma-2020>
4. GitHub - som-shahlab/Clinfo.AI: This is Clinfo.AI Demo Instruction [Internet]. [cited 2024 Jul 29]. Available from: <https://github.com/som-shahlab/Clinfo.AI>
5. Semantic Scholar | AI-Powered Research Tool [Internet]. [cited 2024 Jul 31]. Available from: <https://www.semanticscholar.org/>
6. Lozano A, Fleming SL, Chiang CC, Shah N. Clinfo.ai: An open-source retrieval-augmented large language model system for answering medical questions using scientific literature. Pacific Symposium on Biocomputing [Internet]. 2024 [cited 2024 Jul 29];(2024):8–23. Available from: <https://github.com/som-shahlab/Clinfo.AI>
7. Khan K, Kunz R, Kleijnen J, Antes G. Systematic reviews to support evidence-based medicine, 2nd edition. Systematic reviews to support evidence-based medicine, 2nd edition. 2011.
8. Chiari P, Mosci D, Naldi E. Evidence-based clinical practice la pratica clinico-assistenziale basata su prove di efficacia. 2. ed. McGraw-Hill, editor. Milano; 2011.