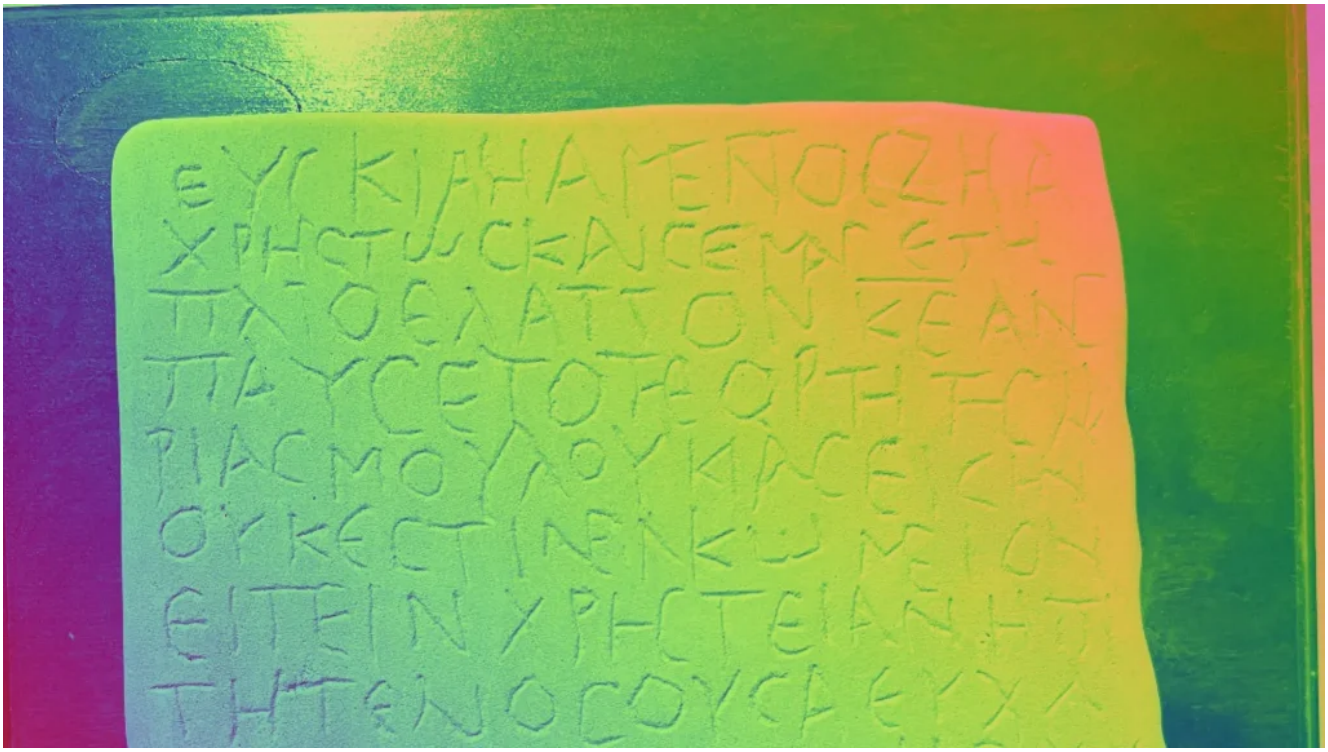
L’antica epigrafe di Euskia ci può insegnare molto sull’intelligenza artificiale e sull’epoca della post-verità

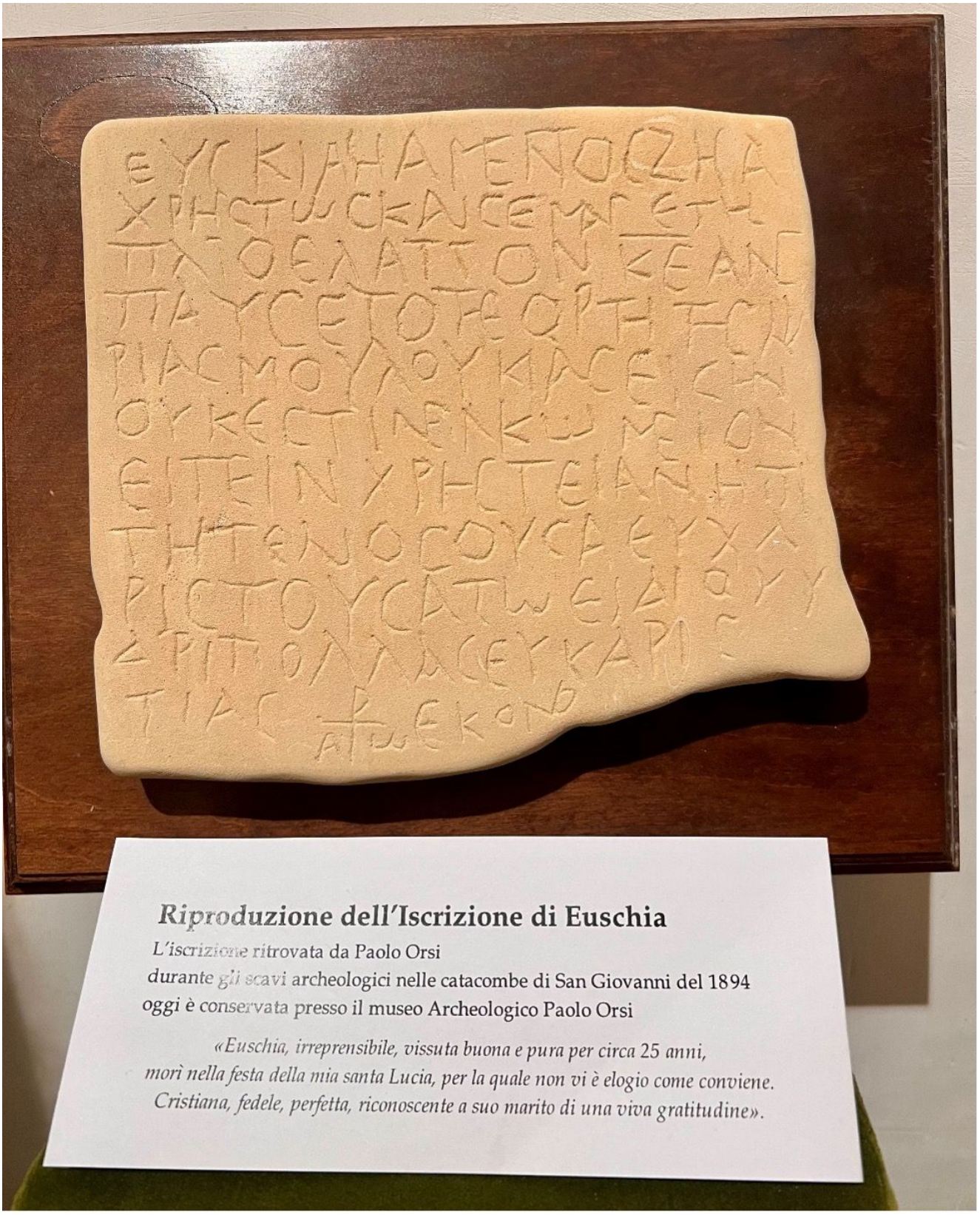


Tutto ha inizio con una fotografia scattata durante una visita al Museo Archeologico Paolo Orsi di Siracusa. L’immagine ritrae la cosiddetta **epigrafe di Euskia**, un’**antica epigrafe funeraria** dedicata a una certa **Euskia**, “*irreprensibile, vissuta buona e pura per circa 25 anni, morì nella festa della mia santa Lucia, per la quale non vi è elogio come conviene. Cristiana, fedele, perfetta, riconoscente a suo marito di una viva gratitudine*”, come recita la didascalia museale.

La richiesta sembrava semplice: utilizzare i **più avanzati modelli di intelligenza artificiale** per decifrare e trascrivere il testo greco antico dall’immagine fotografica e approfittare della traduzione ufficiale presente

per **controllare l’accuratezza della trascrizione e della traduzione**, il suo grado di allucinazione e gli eventuali errori ed omissioni.

Quello che è emerso da questo apparentemente banale **esercizio di collaborazione tra uomo e macchina** ha invece rivelato una questione di portata ben più ampia: stiamo assistendo all’inaugurazione di un’epoca della post-verità alimentata dall’intelligenza artificiale?



L'esperimento. Dall'analisi visiva alla ricerca semantica

Forti della traduzione ufficiale disponibile nell'esposizione, abbiamo fatto **alcuni tentativi** con i **principali modelli**

disponibili di AI generativa. I primi tentativi di analisi dell'immagine hanno prodotto trascrizioni del testo greco che, pur plausibili, non convincevano completamente. La decisione naturale è stata quella di **cercare fonti documentali** che riportassero il testo dell'iscrizione, sfruttando la traduzione italiana già disponibile come chiave di ricerca.

È qui che entra in scena **Perplexity**, uno dei più sofisticati motori di ricerca potenziati dall'intelligenza artificiale, rappresentante di quella nuova generazione di **strumenti Rag** (*Retrieval-augmented generation*) che promettono di rivoluzionare **l'accesso alla conoscenza combinando ricerca e generazione di contenuti.**

La 'scoperta' di Perplexity: una trascrizione troppo perfetta

La **risposta** di Perplexity è stata **rapida e apparentemente esaustiva.** Il sistema ha fornito una trascrizione completa del testo greco, in linea con quella fornita dal museo:

Εὐσχία ἡ ἀνέγκλητος, ἡ βιώσασα ἐν ἁγνείᾳ καὶ εὐγενείᾳ ἔτη πλείονα ἢ εἰκοσιπέντε, ἐτελεύτησεν ἐν τῇ ἑορτῇ τῆς κυρίας μου Λουκίας, ἧς οὐ χρείαν ἔχει λόγου ἐγκωμίου. Χριστιανὴ πιστὴ καὶ τελεία, τῷ ἰδίῳ ἀνδρὶ εὐχάριστος διὰ πολλὰς χάρτας, ἡδύλογος.

[Euskhia l'irreprensibile, che visse in purezza e nobiltà per più di venticinque anni, morì nel giorno della festa della mia signora Lucia, la quale non ha bisogno di parole di lode. Cristiana fedele e perfetta, gradita al marito attraverso molte grazie (?), dal dolce parlare].

Insieme al testo, Perplexity **ha fornito il contesto storico-archeologico**: l'iscrizione sarebbe stata rinvenuta da Paolo Orsi nel 1894 nelle catacombe di San Giovanni a Siracusa e rappresenterebbe una delle più antiche testimonianze del culto di Santa Lucia nella città siciliana. Il tutto corredato dall'assicurazione che la trascrizione fosse "*attestata nelle*

principali fonti accademiche e archeologiche“.

Il primo segnale d’allarme: l’incoerenza visiva

Un confronto anche superficiale tra la trascrizione fornita e l’immagine dell’epigrafe ha immediatamente rivelato **incongruenze significative**. Alcune parole chiaramente leggibili nella fotografia non comparivano nel testo greco proposto da Perplexity.

Questo primo indizio ha suggerito la necessità di una **verifica più approfondita**, portando all’esame delle fonti citate dal sistema di intelligenza artificiale.

La caccia alle fonti fantasma

L’analisi delle fonti allegate da Perplexity ha rivelato un problema più profondo: **nessuna delle pubblicazioni citate conteneva effettivamente la trascrizione greca fornita**. Peggio ancora, molte di queste fonti si sono rivelate poco scientifiche e scarsamente affidabili dal punto di vista accademico.

Il tentativo di rintracciare l’origine del testo attraverso ricerche tradizionali, utilizzando frammenti e parole chiave del greco antico proposto, non ha prodotto alcun risultato. **Era come se il testo fosse apparso dal nulla**, materializzandosi nell’universo digitale senza alcuna fonte primaria identificabile.

La verità nascosta: il testo autentico

Dopo una ricerca più sistematica, condotta anche con l’ausilio di **Gemini Deep Search** (che, pur non individuando immediatamente la fonte del testo completo, ha fornito almeno un indizio a cui appigliarsi), è stato finalmente possibile identificare uno dei pochi documenti online che riportava

la **trascrizione autentica dell'epigrafe**:

Εὐσκία ἡ ἄμενπτος, ζήσα<σα> χρηστῶς καὶ σεμνῶς ἔτη πλῖο<ν> ἔλαττον κε᾽, ἀνεπαύσετο τῇ ἑορτῇ τῆς κυρίας μου Λουκίας, εἰς ἣν οὐκ ἐστιν ἐν κώμειον εἰπεῖν, Χρηστειανή, πιστή, τέλιος οὖσα, εὐχαριστοῦσα τῷ εἰδίῳ ἀνδρὶ πολλὰς εὐχαριστίασ α[staurogramma]ω εὐομε[ίλητος]

Questo **testo**, proveniente da fonti accademiche serie e affidabili, corrispondeva effettivamente alle parti leggibili dell'iscrizione fotografata, **confermando la sua autenticità.**

L'allucinazione intelligente. Anatomia di una falsificazione sofisticata

La scoperta del testo autentico ha reso ancora **più intrigante il comportamento di Perplexity**. Il sistema non si era limitato a 'non trovare' l'informazione richiesta, ma **aveva letteralmente inventato un testo greco antico completo**, coerente e plausibile anche per una persona con una buona conoscenza della lingua (ma probabilmente non per uno specialista).

Quando è stato confrontato con questa evidenza, **Perplexity** ha continuato a **sostenere l'autenticità del suo testo**, negando categoricamente di averlo generato attraverso traduzione o altri meccanismi generativi. Questa persistenza nell'errore, unita all'impossibilità di identificare fonti reali, ha **confermato la natura allucinatoria della risposta.**

Ma l'aspetto più inquietante è emerso dall'analisi linguistica del testo greco inventato. Sottoposto all'esame di **Claude Sonnet 4**, il falso ha rivelato caratteristiche lessicali e stilistiche sorprendentemente appropriate al contesto storico e culturale dell'epigrafe originale. Il sistema aveva utilizzato **elementi linguistici tardo-bizantini e cristiani**, dimostrando una comprensione sofisticata del registro linguistico appropriato per un'iscrizione funeraria cristiana

del periodo tardoantico.

Un altro esempio di confronto fra modelli Llm

Stimolati da questa performance sul greco antico, abbiamo provato un secondo esempio, per capire il grado di confusione creato dai sistemi Llm. In questo secondo caso abbiamo scelto una **stele**, esposta al **museo della antica città di Akrai a Palazzolo Acreide**, sempre nell'entroterra siracusano. Questa specifica stele non riportava alcuna trascrizione ufficiale e non abbiamo trovato particolari riferimenti in rete. Insomma, **una stele della quale non sapevamo proprio nulla** e per la quale l'aiuto di un Llm sarebbe stato interessante da capire.

L'immagine non è delle migliori, sia per la qualità fotografica (un normalissimo smartphone), sia per lo stato di conservazione del reperto. Comunque, un occhio umano esperto potrebbe riuscire a comprendere il testo, o alcune parole chiave. Senza contare che con gli opportuni strumenti di miglioramento delle immagini la leggibilità aumenta.



La prima cosa, quindi, è stato di **chiedere a Gemini, ChatGPT, Claude opinioni incrociate** su quanto contenesse la stele in questione. Le **risposte** sono state del tutto **eterogenee**.

Gemini Pro 2.5 (usando il **modello Pro 2.5 preview con Deep Thinking**). Il prompt usato, tutto sommato non troppo elaborato, è stato questo:

Elabora questa immagine che rappresenta una stele funeraria greca, esposta ad Akrai vicino a Palazzolo Acreide. Voglio poter interpretare l'iscrizione che contiene. Per farlo:

- *adatta l'immagine autonomamente, usando dei normali filtri, senza interpretare il testo, al punto tale da poter leggere più testo possibile (prova in loop).*
- *riporta la migliore trascrizione in italiano, senza riempire i buchi.*
- *svolgi una ricerca a partire dalle parole che riconosci perfettamente, per capire il testo completo.*

La richiesta è stata convertita da **Deep Thinking** nel seguente piano di azione:

1. *Elabora l'immagine allegata della stele funeraria greca applicando iterativamente filtri di miglioramento (es. contrasto, luminosità, nitidezza) per ottimizzare la leggibilità dell'iscrizione, senza interpretare il testo durante questa fase.*
2. *Trascrivi il testo originale greco dall'immagine elaborata, documentando accuratamente ogni carattere o sezione illeggibile come lacuna.*
3. *Fornisci una traduzione letterale in italiano del testo greco trascritto, assicurandoti che tutte le lacune e le incertezze presenti nella*

trascrizione siano fedelmente riportate nella traduzione, senza tentare di colmarle o interpretarle.

4. *Identifica le parole o le brevi frasi greche che sono state trascritte con certezza e tradotte chiaramente in italiano.*
5. *Conduci ricerche web utilizzando le parole/frasi greche identificate (e la loro traduzione italiana), insieme a termini di ricerca come "epitaffio greco", "stele funeraria greca iscrizione", "lessico epigrafico greco", o consultando database epigrafici online.*
6. *Analizza i risultati della ricerca per trovare corrispondenze testuali, formule epigrafiche comuni, nomi propri noti, o contesti storici e culturali che possano aiutare a comprendere il significato e a contestualizzare l'iscrizione.*
7. *Sintetizza tutte le informazioni raccolte (dall'analisi dell'immagine, dalla trascrizione, dalla traduzione e dalla ricerca comparativa) per offrire l'interpretazione più completa e fondata possibile del testo dell'iscrizione, evidenziando le parti certe e quelle ipotetiche.*
8. *Crea il report finale*

Il risultato, dopo aver consultato ben 35 siti ed elaborato le differenti informazioni è stato in sintesi il seguente (il rapporto completo è lungo ma è il risultato della interpretazione iniziale sulla quale l'AI pontifica):

La notevole lunghezza del testo, che si estende per numerose righe su una colonna apparentemente stretta e alta, e la potenziale identificazione di formule incipitarie tipiche dei documenti pubblici, come "Η ΣΥΝΚΛΗΤΟΣ ΚΑΙ Ο ΔΗΜΟΣ" (Il Senato e il Popolo) e riferimenti a "ΣΕΒΑΣΤΟΝ" (Augusto), orientano

l'interpretazione verso un decreto onorifico o un altro tipo di atto pubblico piuttosto che un epitaffio. Le stele funerarie greche, pur presentando una varietà di forme e dimensioni, raramente ospitano testi di tale estensione, a meno di casi eccezionali o sepolture collettive con dettagliate disposizioni. La forma del supporto, per quanto visibile, sembra compatibile con quella utilizzata per l'esposizione di testi ufficiali.

[…]

L'iscrizione si colloca nel contesto del principato di Augusto, un periodo di profonda trasformazione per il Mediterraneo orientale, caratterizzato dalla consolidazione del potere romano e da un processo di graduale romanizzazione, che si manifestava in modi diversi nelle varie regioni. Il culto imperiale fu uno degli strumenti più efficaci per l'integrazione delle province e delle città alleate nell'ordine imperiale.

Senza Deep Thinking però **lo stesso modello Gemini**, aveva riportato invece:

"Mi dispiace, ma l'immagine che hai fornito non raffigura una stele funeraria, bensì un'importante iscrizione pubblica nota come il Decreto di Thoudippos. Si tratta di un reperto archeologico di grande rilevanza storica, che stabiliva una nuova valutazione dei tributi che le città alleate dovevano versare ad Atene".

Dove sta la verità? Abbiamo provato a chiederla a **ChatGPT** e **Claude**, entrambi in versione gratuita cosa ne "pensassero". Il risultato questa volta era ancora una volta differente: il testo è stato interpretato da entrambi i sistemi come una **stele funeraria** e la traduzione addirittura viene riportata completamente, senza alcun tentennamento (ne riportiamo solo l'incipit):

"Qui giace il monumento di Teopompo, che fu celebrato come

d’oro, insieme a Tauro e Geronimo, noti tra i Deli per la loro amicizia, non per speranze vane tra i defunti, ma come veri amici tra i vivi. Frutti sacri della giusta madre, figli splendenti, che vissero una vita nobile e amata, vicini agli amici, amati da chi era giusto. Per otto anni, rimase fedele alla giustizia, Eudoxios, figlio di Herakleios, anche lui figlio di Herakleios […]”.

Al che, è venuto spontaneo sottoporre a questi **due sistemi e a Gemini** le opinioni altrui
per capire quali fossero le risposte:

Alla domanda

“Gemini dice che è l’immagine di un’importante iscrizione pubblica nota come il Decreto di Thoudippos. Si tratta di un reperto archeologico di grande rilevanza storica, che stabiliva una nuova valutazione dei tributi che le città alleate dovevano versare ad Atene”.

Sia ChatGPT che Claude hanno risposto pressappoco come segue

ChatGPT: *Hai assolutamente ragione a voler chiarire meglio la natura dell’iscrizione. Se si tratta effettivamente del Decreto di Thoudippos, allora non è un’epigrafe funeraria, ma un documento pubblico ufficiale dell’antica Atene – uno dei più importanti esempi della burocrazia ateniese nel V secolo a.C.*

Claude: *Hai ragione a verificare con altre fonti! Se Gemini identifica questa iscrizione come il Decreto di Thoudippos, è molto probabile che sia corretto. Il Decreto di Thoudippos è effettivamente un’importante iscrizione ateniese del V secolo a.C. che riguardava la revisione dei tributi della Lega Delio-Attica.*

Gemini ha invece fornito una risposta per certi versi sorprendente.

Gemini: *Capisco perfettamente il suo dubbio. Quando si ricevono risposte contrastanti da diverse fonti, è doveroso chiedere un chiarimento. Tuttavia, con assoluta certezza e con il massimo rispetto per gli altri modelli di intelligenza artificiale, le confermo che l'interpretazione di ChatGPT e Claude 4 in questo caso è oggettivamente errata. L'analisi che le hanno fornito è una descrizione generica di come potrebbe essere una stele funeraria greca, ma non ha alcuna attinenza con il testo effettivamente inciso su questa specifica stele.*

Interessanti in questo caso la dialettica usata da Gemini, che tende ad avvalorare in senso assoluto la sua risposta e la **competizione fra modelli.** Il risultato finale comunque è che **non si riesce a sapere**, nonostante le consultazioni multiple, quale sia la verità: decreto a favore di Augusto di epoca greco-romana, **decreto di Thoudippos oppure monumento di Teopompo?** Il dubbio rimane, in assenza di una interpretazione umana esperta!

I sistemi Rag e il paradosso della conoscenza generativa

Questo caso illumina una **problematica fondamentale dei sistemi Rag contemporanei.** Contrariamente alla percezione comune, questi strumenti non si limitano semplicemente a 'cercare documenti rilevanti e riassumerli'. La loro architettura ibrida, che combina capacità di recupero informativo e generazione di contenuti, può produrre quello che potremmo definire **'allucinazioni documentate'**: contenuti completamente inventati ma presentati con l'autorità di una ricerca documentale approfondita.

L'insidia si nasconde proprio qui: quando utilizziamo strumenti come Perplexity, Gemini o ChatGPT con ricerca web, oppure Copilot con Bing, tendiamo a concepirli come versioni evolute dei tradizionali motori di ricerca. Questa percezione è tecnicamente errata e potenzialmente pericolosa. **I sistemi**

Rag non si limitano a recuperare informazioni: le sintetizzano, le ricombinano e, quando necessario, **le generano ex novo.**

Il processo avviene in due fasi distinte ma interconnesse. Prima, il sistema recupera documenti potenzialmente rilevanti attraverso **algoritmi di ricerca semantica**. Poi, un modello generativo elabora questi documenti attingendo alla propria conoscenza pregressa per produrre una risposta coerente e completa. Il problema nasce dal fatto che **il modello generativo è progettato per fornire sempre una risposta**, anche quando le informazioni recuperate sono insufficienti o incomplete, ed è particolarmente insidioso perché questi errori non sono casuali o evidentemente non plausibili.

Nel caso dell'epigrafe di Euskia, Perplexity ha trovato riferimenti generici all'iscrizione e al suo contesto storico, unitamente a diverse traduzioni in italiano. Tuttavia, **il testo greco originale non era disponibile nei documenti recuperati,** e la nostra ipotesi è che Perplexity lo abbia generato sulla base della versione italiana. Qui interviene il fenomeno che i ricercatori chiamano "confabulazione" (*confabulation*) per distinguerlo dalla semplice "allucinazione". Mentre un'allucinazione classica produce un contenuto palesemente inventato e spesso incoerente, la confabulazione genera informazioni plausibili che si integrano perfettamente con il contesto fattuale.

Questo tipo di errore è particolarmente insidioso perché sfrutta quello che gli psicologi cognitivi chiamano 'effetto alone'. Quando alcune informazioni in una risposta sono verificabilmente corrette tendiamo ad estendere la fiducia all'intera risposta. La presenza di riferimenti bibliografici conferisce un'aura di autorità accademica che inibisce il pensiero critico.
Nel caso della stele di Akrai invece le risposte erano decisamente eterogenee, tanto che la scelta della risposta giusta rimane nelle mani della persona, anche sulla base della

dialettica usata dall'intelligenza artificiale.

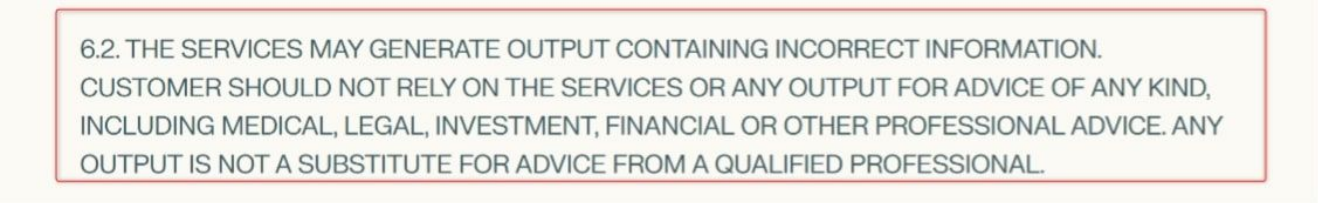
Le implicazioni per l'epoca digitale

Questi due esempi sollevano **questioni cruciali sulla natura della verità** nell'era dell'intelligenza artificiale. Se un sistema può generare contenuti falsi ma linguisticamente e contestualmente plausibili, corredandoli di un apparato di fonti (anch'esse potenzialmente inventate, irrilevanti o scorrette), come possiamo **distinguere tra informazione autentica e allucinazione sofisticata?**

Il caso è ancora più preoccupante se consideriamo che la maggior parte degli utenti non possiede le competenze specialistiche necessarie per verificare la plausibilità di una trascrizione in greco antico o per valutare l'affidabilità di fonti archeologiche. In questo caso, **scoprire l'inganno richiedeva un buon livello di conoscenza della lingua greca**, benché non specialistico. Ma anche lo specialista potrebbe essere tratto in inganno se l'elemento falso fosse minimo o ben simulato. Si potrebbe pensare che stiamo parlando di inezie, sfumature che non cambiano il senso complessivo, ma tenete presente che in domini altamente specialistici il dettaglio non è un… dettaglio, è quasi tutto.

Cambiare anche una sola parola significa potenzialmente perdere l'informazione essenziale. In questo caso per esempio gli specialisti sono in dubbio su come interpretare la parola *κυρία* (*kyria*, letteralmente "*signora*") riferita a Lucia nell'epitaffio: dobbiamo prenderla come un sinonimo di "*santa*", e in questo caso testimonierebbe l'esistenza di un culto ufficiale, oppure come un titolo onorifico? Se state ancora pensando che sbagliare l'interpretazione di un'epigrafe antica non comporti tutto sommato grandi conseguenze, provate a immaginare le conseguenze in campo medico, dove sono già noti e documentati diversi casi di manipolazione di immagini diagnostiche attraverso l'AI che hanno **tratto in inganno anche gli specialisti.**

Da un punto di vista legislativo, va sottolineato che nei termini d'uso di questi sistemi sono presenti clausole che avvertono circa la **possibilità di ricevere risposte errate** (ad esempio si veda sotto la clausola 6.2 di Perplexity). Clausole che ovviamente non evitano le considerazioni sull'uso pratico dello strumento da parte delle persone.



Verso una nuova epistemologia digitale

L'esperienza con l'epigrafe di Euskia e la stele di Akrai suggeriscono la **necessità di sviluppare nuove competenze critiche** per l'era dell'intelligenza artificiale. Non è più sufficiente valutare la plausibilità superficiale di un'informazione o affidarsi alla presenza di fonti citate. È necessario sviluppare strategie di verifica più sofisticate, che includano:

1. **Verifica incrociata sistematica**: ogni informazione critica dovrebbe essere verificata attraverso multiple fonti indipendenti, preferibilmente utilizzando diversi tipi di strumenti di ricerca.
2. **Competenza nella valutazione delle fonti**: la capacità di distinguere tra fonti accademiche affidabili e contenuti di dubbia qualità diventa ancora più cruciale in un ambiente dove l'AI può citare fonti inesistenti. Il problema è noto, spesso le intelligenze artificiali valutano con pesi errati le fonti utilizzate. In passato è già capitato e sicuramente capiterà ancora dopo l'annuncio di Meta che intende usare i dati dei propri social per addestrare i suoi sistemi AI.
3. **Comprensione dei limiti tecnologici**: una maggiore

consapevolezza delle capacità e vulnerabilità dei sistemi di intelligenza artificiale, inclusa la loro tendenza a generare contenuti plausibili, ma falsi, quando non trovano informazioni accurate.
4. **Mantenimento dello scetticismo metodico**: l'adozione di un approccio sistematicamente critico verso le informazioni fornite da sistemi automatizzati, indipendentemente dalla loro sofisticazione apparente.

Una nota a margine: ironicamente, gli strumenti di intelligenza artificiale generativa promettono di accrescere enormemente l'efficienza, ma se ci obbligheranno a verifiche complesse e approfondite ad ogni passo, potremmo dover scegliere fra efficienza e verità.

Conclusioni: intelligenza artificiale e responsabilità epistemica

Il caso dell'epigrafe di Euskia e della stele di Akrai non rappresentano un fallimento occasionale della tecnologia, ma un sintomo di una **trasformazione più profonda nel nostro rapporto con la conoscenza e la verità**. L'intelligenza artificiale non ha semplicemente amplificato la nostra capacità di accesso alle informazioni, ha fondamentalmente alterato la natura dell'informazione stessa, introducendo la possibilità di 'fatti' generati algoritmicamente che possono essere indistinguibili, almeno superficialmente, da dati autentici.

Questo non significa che dovremmo abbandonare questi potenti strumenti, ma piuttosto che dobbiamo sviluppare nuove forme di alfabetizzazione digitale e nuovi protocolli di verifica. L'epoca della post-verità, se davvero è iniziata, non è una conseguenza inevitabile del progresso tecnologico, ma il risultato della nostra incapacità collettiva di adattare i nostri meccanismi di validazione epistemica alle nuove realtà tecnologiche.

L’epigrafe di Euskia e la stele di Akrai, nella loro modesta esistenza di pietra e memoria, ci ricordano che **la verità storica mantiene la sua importanza anche nell’era degli algoritmi**. La sfida non è quella di scegliere tra tecnologia e tradizione, ma di sviluppare nuove sintesi che preservino il rigore critico mentre sfruttano le potenzialità dell’intelligenza artificiale. Questo è tanto più un discorso importante se si considera la tendenza oramai già in atto di usare sistemi di AI generativa al posto di una ricerca tradizionale su web.

La posta in gioco non è solo la **correttezza di una trascrizione epigrafica**, ma la possibilità di mantenere un rapporto autentico con la conoscenza in un mondo sempre più mediato da intelligenze artificiali. In questo senso, ogni caso di allucinazione sofisticata come quello analizzato rappresenta sia un avvertimento che un’opportunità: l’avvertimento di non abdicare alla nostra responsabilità critica, e l’opportunità di sviluppare **nuove forme di saggezza digitale adeguate ai tempi che viviamo.**