SAPIENZA UNIVERSITÀ DI ROMA

COGNITIVE PHILOLOGY
No 18 (2025)
ISSN 2035-391X

# L'Evoluzione dell'Attribuzione d'Autore nell'Era dei Large Language Models: Tokenizzazione ed Embedding

Mirko Degli Esposti*

*Dipartimento di Fisica e Astronomia, Alma Mater Studiorum - Università di Bologna mirko.degliesposti@unibo.it

"Our reality isn't just described by mathematics – it is mathematics… Not just aspects of it, but all of it, including you. In other words, our external physical reality is a mathematical structure." (Max Tegmark, *Our Mathematical Universe*, Chapter 12)

## 1. Premessa

In qualità di fisico matematico che si è occupato a lungo di attribuzione d'autore attraverso metodi entropici, basati su compressori e distanze di n-grammi (Mekala et al. 2018; Basile – Benedetto – Caglioti – Degli Esposti 2008), è per me evidente che la rivoluzione dei Large Language Models (LLM) ha profondamente modificato il panorama della analisi, della comprensione e della generazione artificiale di testi naturali (Brown et al. 2020). Questa trasformazione ha innescato nuove sfide e opportunità, rendendo, a mio avviso, il campo della filologia attributiva ancor più stimolante. Le ripercussioni di questi modelli sono già visibili e stanno ridefinendo velocemente le frontiere della disciplina (Silva et al. 2024; Huang – Chen – Shu 2024).

Per questo breve intervento alla conferenza "Filologia Attributiva - Metodi Computazionali", ho scelto di concentrarmi su due aspetti tecnici e matematici dell'AI generativa che rappresentano in qualche modo un'estensione naturale degli approcci tradizionali all'attribuzione d'autore: la *tokenizzazione* e gli

*embedding*. Questi concetti, pur non essendo interamente nuovi, assumono un ruolo centrale nell'analisi dei testi alla luce delle capacità predittive e generative dei LLM.

## 2. Tokenizzazione: la scomposizione del testo

La tokenizzazione è il processo di suddivisione di un testo in unità discrete, chiamate *token* (Manning – Schütze 1999). Tradizionalmente, i token potevano essere parole, caratteri o sottocategorie di parole. Tuttavia, nell'ambito dei LLM, la tokenizzazione ha assunto una granularità più sofisticata, spesso basata su *subword units*, ossia *unità sub-parola*. Questo approccio ibrido, che combina i vantaggi della tokenizzazione a livello di carattere e di parola, consente di gestire efficacemente parole rare o sconosciute e di ridurre la dimensione del vocabolario, migliorando la generalizzabilità dei modelli (Sennrich et al. 2016; Roy 2025; Sanders 2022). Ma forse è utile procedere con un esempio concreto. Consideriamo ad esempio, il testo di Italo Calvino, *La Città Ragnatela*:

> Se volete credermi, bene, Ora dirò come è fatta Ottavia, città - ragnatela. C'è un precipizio in mezzo a due montagne scoscese: la città è il suo vuoto, legata alle due creste con funi e catene e passerelle. Si cammina sulle traversine di legno, attenti a non mettere il piede negli intervalli, o ci si aggrappa alle maglie di canapa. Sotto non c'è niente per centinaia e centinaia di metri: qualche nuvola scorre; s'intravede più in basso il fondo del burrone. Questa è la base della città: una rete che serve da passaggio e da sostegno. Tutto il resto, invece d'elevarsi sopra, sta appeso sotto: scale di corda, amache, case fatte a sacco, attaccapanni, terrazzi come navicelle, otri d'acqua, becchi del gas, girarrosti, cesti appesi a spaghi, montacarichi, docce, trapezi e anelli per i giochi, teleferiche, lampadari, vasi con piante dal fogliame pendulo. Sospesa sull'abisso, la vita degli abitanti d'Ottavia è meno incerta che in altre città. Sanno che più di tanto la rete non regge.

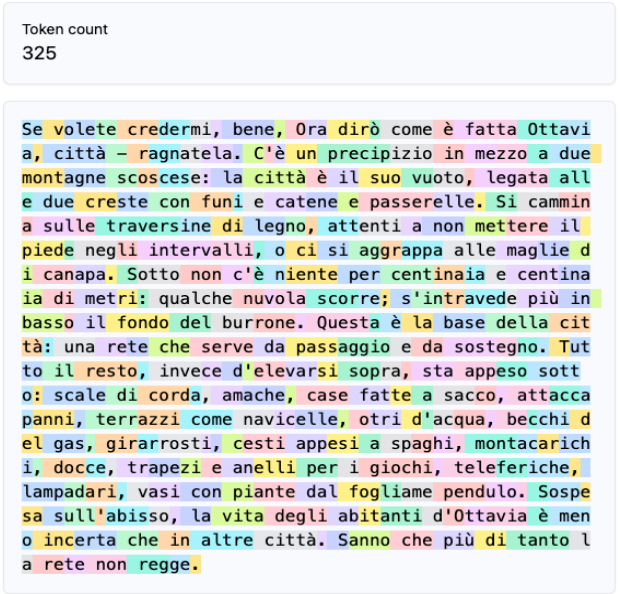


Figura 1 (fonte: https://tiktokenizer.vercel.app/?model=cl100k_base)

È per noi umani naturale interpretare il testo come una successione di parole, oppure al più come una

sequenza di caratteri (spazio incluso). I modelli fondamentali del linguaggio invece 'leggono' (e generano) il testo in maniera piuttosto diversa: come sequenze di tokens appunto, dove un singolo token è una sequenza di caratteri contigui, spesso una parte di una parola completa. Nella figura qui sotto, ad esempio, mostriamo come ChatGPT4 *'tokenizza'* il testo di Calvino: ogni singolo colore rappresenta un token distinto.

- Si noti come ciascun token raramente rappresenta una intera parola ma solo una parte di essa: ad esempio *'ragnatela'* viene letta come successione di tre tokens *' r'*, *'agn'*, *'at'*, *'ela'*, mentre *'nuvola'* è rappresentata da un singolo token che coincide con la parola.

In pratica, algoritmi come *Byte Pair Encoding* (BPE) (Sennrich et al. 2016), *WordPiece* (Devlin et al. 2019) o *SentencePiece* (Kudo – Richardson 2018), costruiscono un vocabolario di token iterativamente, fondendo/comprimendo le coppie di byte o caratteri più frequenti. La scelta del *tokenizzatore* e la sua configurazione – ad esempio, la dimensione del vocabolario – influenzano direttamente la rappresentazione numerica del testo e, di conseguenza, le prestazioni nei compiti di attribuzione. Per inquadrare gli ordini di grandezza, i tokenizzatori impiegati nei modelli di linguaggio più diffusi oscillano fra poche decine di migliaia e oltre centomila unità lessicali distinte. Ad esempio:

- *BERT Base/Large* usa WordPiece con circa 30 000 token, integrati da simboli speciali di controllo;
- *T5* impiega SentencePiece con 32 128 token;
- *Llama 2* adotta un ibrido BPE/SentencePiece intorno a 32 000 token;
- *GPT-3*, attraverso la libreria *tiktoken*, estende il vocabolario a circa 50 000 token (p50k_base), mentre *GPT-4* supera i 100 000 (cl100k_base);
- la recente *Llama 3* introduce un tokenizzatore ulteriormente ottimizzato che tocca 128 000 token, migliorando la copertura multilingue.

Il confronto con il lessico umano mette in rilievo la natura compressiva di queste rappresentazioni: dizionari monumentali come l'*Oxford English Dictionary* catalogano oltre mezzo milione di lemmi e, includendo le forme flessionali, alcune stime superano i due milioni di parole. Ciononostante, un adulto madrelingua inglese conosce attivamente circa 20 000–35 000 parole – fino a 60 000 in riconoscimento passivo. I tokenizzatori, codificando in modo compatto sequenze di caratteri ricorrenti, riescono dunque a rappresentare spazi lessicali enormi con vocabolari relativamente contenuti, offrendo un compromesso fra granularità linguistica ed efficienza computazionale nei compiti di attribuzione. Il confronto è sorprendente:

- I vocabolari dei LLM, pur essendo nell'ordine delle decine o poche centinaia di migliaia di token, sono molto più piccoli del numero potenziale di tutte le forme flesse e derivate di parole in inglese.
- Questa apparente discrepanza è risolta dall'uso dei *subword tokens* – unità *sub-parola*. Invece di memorizzare ogni forma flessa di ogni parola (es. "correre", "corro", "corri", "correva", "corremmo" come token separati), i LLM scompongono queste parole in unità più piccole (es. "corr", "ere", "o", "eva", "emmo"). Questo permette al modello di costruire qualsiasi parola da un insieme limitato di token base, gestendo l'illimitata creatività del linguaggio con un vocabolario gestibile.
- Questa metodologia è cruciale per la generalizzazione e la robustezza dei LLM. Essi non hanno bisogno di aver "visto" ogni singola parola flessa durante l'addestramento per comprenderla o generarla; possono assemblarla combinando diversi tokens.
- In termini di filologia attributiva, ciò significa che l'analisi non si basa più solo su conteggi di parole o n-grammi di caratteri/parole, ma sulla distribuzione e combinazione di queste unità sub-parola e sulla loro rappresentazione nello spazio degli embedding. Lo stile di un autore può manifestarsi non solo nella scelta delle parole, ma anche nella frequenza con cui certe sequenze di token appaiono, o nel modo in cui il tokenizzatore stesso "spezza" le parole di un autore rispetto a un altro.

## 3. Embedding: la mappatura dei token in spazi vettoriali

Una volta segmentato il testo, dobbiamo trasformare ciascun token in un vettore numerico che ne codifichi le proprietà distribuzionali, statistiche, sintattiche e semantiche. Un *embedding* proietta ciascun token in un punto di uno spazio vettoriale di dimensione "d" in modo che la distanza fra i vettori rifletta la vicinanza linguistica fra i corrispondenti concetti testuali (Turney – Pantel 2010; Pennington – Socher – Manning 2014; Alammar 2019). In questo spazio, le parole che compaiono in contesti analoghi – per esempio *gatto* e *felino* – finiscono in regioni contigue e danno luogo a regolarità geometriche che consentono anche operazioni algebriche intuitive – sottrarre *re* e aggiungere *donna* conduce naturalmente verso *regina* (Mikolov – Yih – Zweig 2013). La struttura non è solo semantica: trasporta anche relazioni sintattiche, perché tempi verbali, plurali o varianti prefissate si dispongono lungo direzioni quasi parallele; incorpora dimensioni pragmatiche, dato che lo stile formale e quello colloquiale si distribuiscono su assi latenti distinti; e, nei modelli multimodali, accoglie perfino le corrispondenze fra testo e immagini all'interno di un unico spazio condiviso. Tutte queste proprietà emergono durante il pre-training mediante ad esempio Transformer: ottimizzando l'obiettivo di un

modello di linguaggio, il modello avvicina i vettori associati a contesti simili e separa quelli che ricorrono in ambienti diversi (Vaswani et al. 2017).

La ricchezza di sfumature che l'*embedder* può codificare dipende dalla dimensione d del vettore. Nei modelli più noti questa dimensione varia da qualche centinaio a oltre sedicimila componenti: *BERT-Base* adopera 768 dimensioni, mentre *BERT-Large* si spinge a 1 024; *GPT-2 Small* condivide la scelta di 768, ma *GPT-3*, nel suo formato da 175 miliardi di parametri, amplia lo spazio fino a circa 12 288 dimensioni e *GPT-4*, secondo stime pubbliche, raggiunge quota 16 384. Anche la famiglia *Llama* segue questa crescita: la versione 2 da 7 miliardi di parametri utilizza 4 096 dimensioni, che raddoppiano a 8 192 nel modello con 70 miliardi di parametri, mentre *Llama 3* mantiene 4 096 dimensioni pur introducendo un tokenizzatore più efficiente; L'aumento della dimensionalità offre al modello la possibilità di racchiudere sfumature semantiche sempre più fini, ma impone costi computazionali più elevati sia in fase di addestramento sia di inferenza.

Per confrontare stili d'autore non è necessario analizzare il comportamento di ogni singolo token in isolamento. Una prassi ormai diffusa consiste nell'aggregare i vettori di embedding di una frase o di un paragrafo, trasformandoli in una sorta di *firma latente del testo*. Le tecniche di aggregazione variano dalla media semplice alla somma ponderata per frequenza, fino a metodi più sofisticati come il *pooling guidato dall'attenzione*, in cui il modello stesso determina quali parole o frasi "pesano" di più nella rappresentazione complessiva. Il risultato è un vettore che non rappresenta più un singolo token, ma un'intera unità semantico-stilistica – come una frase, una strofa o una pagina (Le – Mikolov 2014; Lin et al. 2017; Cer et al. 2018).

Queste *firme vettoriali* possono poi essere proiettate in spazi a bassa dimensionalità, attraverso tecniche di riduzione come *PCA* (*Principal Component Analysis*) o *SVD* (*Singular Value Decomposition*)**.** Questo passaggio, oltre a semplificare la visualizzazione e l'analisi, esalta le dimensioni latenti maggiormente rilevanti per la discriminazione stilistica**,** attenuando il rumore e valorizzando i tratti distintivi. In molti esperimenti, queste proiezioni mostrano una separazione marcata tra autori differenti**,** anche nel caso di testi brevi o stilisticamente ambigui.

A differenza dei tradizionali metodi basati su frequenze di n-grammi, che trattano il testo come un oggetto puramente statistico, la rappresentazione tramite embedding incorpora informazioni distribuzionali, sintattiche, pragmatiche e persino stilistiche. Due testi che utilizzano parole diverse ma che trasmettono la stessa struttura logica o retorica possono risultare molto simili nel loro embedding. Al contrario, due testi con parole simili ma stili distinti – per esempio, uno formale e uno colloquiale – si dispongono in regioni lontane dello spazio latente.

Da un punto di vista computazionale, questa metodologia offre un vantaggio notevole: *lo stile emerge*

*come proprietà geometrica*, osservabile e quantificabile attraverso operazioni lineari, distanze e clustering. In altre parole, la distanza fra due firme latenti non è più solo una misura astratta, ma un indicatore concreto di vicinanza stilistica o autoriale. Per questo, l'uso degli embedding aggregati e proiettati rappresenta oggi una delle strategie più efficaci – e al tempo stesso più interpretabili – per affrontare il problema dell'attribuzione d'autore in ambiente LLM.

## 4. Style Transfer e le sue Implicazioni

Una delle aree più suggestive emerse all'intersezione tra linguistica computazionale e generazione automatica di testi è quella dello *style transfer*, ossia la capacità di modificare il registro stilistico di un testo mantenendone inalterato il contenuto semantico. Questo compito, estremamente complesso per un essere umano, è oggi affrontabile dai Large Language Models grazie alla loro rappresentazione latente dei testi sotto forma di embedding separabili: uno legato al contenuto, l'altro allo stile (Jin et al. 2022; Fu et al. 2018; Yang – Carpuat 2025).

La dinamica alla base è, in apparenza, semplice: si prende un testo T, si identifica il suo contenuto latente C(T) e lo si "ri-esprime" secondo uno stile target S_target, ottenendo così un nuovo testo T' che soddisfa due condizioni:

- C(T') ≈ C(T) (il contenuto è conservato)
- S(T') ≈ S_target (lo stile è stato trasformato).

Questa formulazione, che si presta naturalmente a un'interpretazione come problema di ottimizzazione in uno spazio vettoriale, ha ricadute immediate nella filologia attributiva. Trasferire uno stile su un testo di autore ignoto e osservare quanto "bene" o "male" il testo si adatta alla nuova veste può aiutarci a mettere in luce invarianti stilistiche profonde, ossia quei tratti che resistono alla trasformazione e che caratterizzano in modo robusto la mano autoriale.

Per illustrare più concretamente l'idea, consideriamo un semplice esperimento, senza pretese di utilità o rigore scientifico, in cui il brano già usato di Italo Calvino, *La città ragnatela*, è stato riscritto in quattro registri stilistici diversi usando un LLM (GPT3 in questo caso), come riportati in *Appendice*:

1. *Dante Alighieri* – La riscrittura in terzine mostra un lessico solenne, una sintassi incatenata e l'uso di immagini religiose e verticali. L'operazione di style transfer in questo caso ha trasformato non solo il vocabolario, ma anche la struttura metrica e l'intonazione moraleggiante del testo. Il contenuto della città sospesa resta, ma il lettore percepisce un mutamento radicale nell'intenzione espressiva.
2. *James Joyce* – Qui lo stile è iper-letterario e frammentario, con inflessioni da stream of consciousness e uso disinvolto di sintagmi poetici. Il risultato è un testo che mantiene la mappa

semantica di Ottavia ma la immerge in un flusso sintattico anomalo, evocando sensazioni più che concetti. Questo suggerisce che lo stile può deformare profondamente la superficie del testo pur preservando il suo scheletro narrativo.

3. *Romanesco*, stile "*610 (sei uno zero)*"[1] – In questa versione umoristica e dialettale, il contenuto semantico è mantenuto, ma lo stile travolge ogni altro aspetto: lessico gergale, sintassi ellittica, ritmo teatrale. Lo stile funziona come una lente di ingrandimento che mette in risalto la dimensione comica e dissacrante del testo originale. È forse il caso più chiaro di come lo style transfer possa essere utilizzato per rivelare elementi nascosti, espandere le potenzialità espressive del testo e persino stravolgere l'apparato retorico mantenendo intatto il messaggio centrale.
4. *Agatha Christie* (italianizzata) – La prosa diventa misurata, logica, quasi investigativa. Lo stile asciutto, con toni da narrazione gialla, infonde al testo una qualità riflessiva e distaccata, spostando l'attenzione sulla meccanica della descrizione più che sul lirismo della visione. Anche in questo caso, lo style transfer suggerisce un cambiamento di tono e prospettiva senza alterare i dati osservativi della città.

L'insieme di questi esperimenti evidenzia che lo style transfer non è solo una curiosità linguistica, ma può essere uno strumento euristico. Permette di esplorare il comportamento del testo attraverso trasformazioni controllate, identificando le componenti stilistiche che più facilmente si adattano e quelle che resistono. Inoltre, offre un nuovo strumento per la comparazione autoriale: invece di confrontare direttamente due testi, possiamo trasformare uno nello stile dell'altro e misurare quanto l'operazione riesca o fallisca, valutando così la distanza stilistica implicita.

In prospettiva, lo style transfer potrebbe diventare una tecnica ausiliaria per l'attribuzione: in presenza di testi di autore ignoto, si potrebbe cercare lo stile target che minimizza la perdita semantica e massimizza la coerenza retorica, restituendo un indice di affinità autoriale basato su trasformazioni latenti, piuttosto che su confronti superficiali.

## 5. Oltre il Testo: Cenni sulla Multimodalità

Infine, vorrei concludere con un accenno, solo in apparenza marginale, a quella che probabilmente sarà la prossima grande rivoluzione per la filologia computazionale: la *multimodalità*.

I modelli di linguaggio di nuova generazione, come GPT-4V, Gemini o Claude 3, non si limitano più a trattare il linguaggio scritto, ma sono in grado di processare simultaneamente testo, immagini, audio e

[1] https://www.raiplaysound.it/programmi/lilloegreg610.

video all'interno di uno spazio vettoriale condiviso. In altre parole, il concetto stesso di "embedding" si espande oltre il testo, diventando una vera e propria rappresentazione multimodale del mondo (Akkus et al. 2022; OpenAI 2023; Gemini Team Google 2023).

Le implicazioni per l'attribuzione d'autore sono profonde. Si pensi, ad esempio, all'analisi di un manoscritto antico: oltre al contenuto testuale, possiamo ora integrare informazioni visive (tipo di inchiostro, qualità della carta, caratteristiche grafiche e calligrafiche), elementi di impaginazione (struttura della pagina, uso degli spazi, decorazioni marginali), o persino caratteristiche fonetiche nel caso di testi recitati o trascrizioni orali.

Ciascuno di questi canali contribuisce a formare una firma autoriale complessa, stratificata, che difficilmente può essere ridotta alla sola parola scritta. Un modello multimodale può apprendere ad associare il tono vocale a uno stile, a distinguere la mano di uno scriba osservando la pressione o la regolarità del tratto, o ancora a rilevare ricorrenze nella composizione visiva delle opere. Si apre così la possibilità di definire *autorie transmediali,* in cui lo stile non è solo linguistico, ma anche grafico, sonoro, iconografico.

Dal punto di vista tecnico, questi modelli apprendono a mappare simultaneamente input di natura diversa (una descrizione, un'immagine, una voce narrante) in uno stesso spazio semantico. E ciò permette confronti incrociati, ad esempio fra la prosa di un autore e i bozzetti che accompagnano il testo, o fra il ritmo narrativo e il tono vocale di una lettura.

Queste intersezioni fra media possono diventare veri e propri indicatori stilistici, affiancando i tradizionali tratti lessicali o sintattici.

Questa è, forse, la massima estensione del concetto di "spazio vettoriale" che abbiamo fin qui esplorato: non più soltanto spazio delle parole, ma spazio delle espressioni umane – dove testo, voce, immagine e gesto convergono in rappresentazioni dense e navigabili. In questo scenario, la filologia attributiva non è più confinata alla pagina, ma diventa un'esplorazione a tutto campo dell'identità autoriale, nei suoi aspetti linguistici, grafici e performativi.

In sintesi, la tokenizzazione avanzata, la rappresentazione tramite embedding e le capacità trasformative dei LLM, quando arricchite dalla prospettiva della multimodalità, non sono semplicemente estensioni dei nostri vecchi strumenti: rappresentano un cambio di paradigma epistemologico, che ci fornisce strumenti analitici di potenza e granularità senza precedenti per l'attribuzione d'autore nel XXI secolo.

Spero che questo breve excursus possa stimolare nuove riflessioni e direzioni di ricerca nel campo della filologia attributiva attraverso metodi computazionali.

**Riferimenti bibliografici**


Akkus et al. 2022
Cem Akkus – Luyang Chu – Vladana Djakovic – Steffen Jauch-Walser – Philipp Koch – Giacomo Loss – Christopher Marquardt – Marco Moldovan – Nadja Sauter – Maximilian Schneider – Rickmer Schulte – Karol Urbanczyk – Jann Goschenhofer – Christian Heumann – Rasmus Hvingelby – Daniel Schalk – Matthias Aßenmacher, *Multimodal Deep Learning*, «LMU Seminar Series» (web), 30 settembre 2022, URL: https://slds-lmu.github.io/seminar_multimodal_dl/.

Alammar 2019
Jay Alammar, *An Illustrated Guide to Word Embeddings*, «Jay Alammar Blog» (blog), 2019, URL: https://jalammar.github.io/illustrated-word2vec/. jalammar.github.io

Basile – Benedetto – Caglioti – Degli Esposti 2008
Chiara Basile – Dario Benedetto – Emanuele Caglioti – Mirko Degli Esposti, *An Example of Mathematical Authorship Attribution*, «Journal of Mathematical Physics», 49 (2008), 125211, doi: 10.1063/1.2996507.

Brown et al. 2020
Tom B. Brown – Benjamin Mann – Nick Ryder – Melanie Subbiah – Jared Kaplan – Prafulla Dhariwal – Arvind Neelakantan – Pranav Shyam – Girish Sastry – Amanda Askell – Sandhini Agarwal – Ariel Herbert-Voss – Gretchen Krueger – Tom Henighan – Rewon Child – Aditya Ramesh – Daniel M. Ziegler – Jeffrey Wu – Clemens Winter – Christopher Hesse – Mark Chen – Eric Sigler – Mateusz Litwin – Scott Gray – Benjamin Chess – Jack Clark – Christopher Berner – Sam McCandlish – Alec Radford – Ilya Sutskever – Dario Amodei, *Language Models are Few-Shot Learners*, «Advances in Neural Information Processing Systems», 33 (2020), pp. 1877-1901, doi: 10.48550/arXiv.2005.14165. arxi

Cer et al. 2018
Daniel Cer – Yinfei Yang – Sheng-yi Kong – Nan Hua – Nicole Limtiaco – Rhomni St John – Noah Constant – Mario Guajardo-Céspedes – Steve Yuan – Chris Tar – Yun-Hsuan Sung – Brian Strope – Ray Kurzweil, *Universal Sentence Encoder*, Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations,«arXiv preprint arXiv:1803.11175», (2018).

Devlin et al. 2019
Jacob Devlin – Ming-Wei Chang – Kenton Lee – Kristina Toutanova, *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*, «Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)», 1 (2019), pp. 4171-4186, doi: 10.18653/v1/N19-1423. aclantholo

Fu et al. 2018
Zhenxin Fu – Xiaoye Tan – Nanyun Peng – Dongyan Zhao – Rui Yan, *Style Transfer in Text: Exploration and Evaluation*, «Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence», 32 (1) (2018), pp. 663-670, doi: 10.48550/arXiv.1711.06861. cdn.aaai.org

Gemini Team Google 2023

Gemini Team Google, *Gemini: A Family of Multimodal Foundation Models*, «arXiv preprint arXiv:2312.11805», (2023).

Huang – Chen – Shu 2024a
Baixiang Huang – Canyu Chen – Kai Shu, *Authorship Attribution in the Era of LLMs: Problems, Methodologies, and Challenges*, arXiv preprint arXiv:2408.08946, 2024.

Huang – Chen – Shu 2024b
Baixiang Huang – Canyu Chen – Kai Shu, *Can Large Language Models Identify Authorship?*, «Findings of the Association for Computational Linguistics: EMNLP 2024», Miami: Association for Computational Linguistics, 2024, pp. 445-460, doi: 10.18653/v1/2024.findings-emnlp.26.

Jin et al. 2022
Di Jin – Zhijing Jin – Zhiting Hu – Olga Vechtomova – Rada Mihalcea, *Deep Learning for Text Style Transfer: A Survey*, «Computational Linguistics», 48 (1) (2022), pp. 155-205, doi: 10.1162/coli_a_00426. aclanthology.org

OpenAI 2023
OpenAI, *GPT-4V(ision) System Card*, «OpenAI Technical Report», (2023), URL: https://openai.com/research/gpt-4v-system-card

Kudo e Richardson 2018
Taku Kudo e John Richardson, *SentencePiece: A Simple and Language-Independent Subword Tokenizer and Detokenizer for Neural Text Processing*, «Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations», (2018), pp. 66-71, doi: 10.18653/v1/D18-2012.

Le – Mikolov 2014
Quoc V. Le e Tomas Mikolov, *Distributed Representations of Sentences and Documents*, «Proceedings of the 31st International Conference on Machine Learning», 32 (2014), pp. 1188-1196.

Lin et al. 2017
Zhouhan Lin – Minwei Feng – Cicero Nogueira dos Santos – Mo Yu – Bing Xiang – Bowen Zhou – Yoshua Bengio, *A Structured Self-Attentive Sentence Embedding*, «International Conference on Learning Representations (ICLR)», (2017).

Manning – Schütze 1999
Christopher D. Manning e Hinrich Schütze, *Foundations of Statistical Natural Language Processing*, Cambridge (MA): MIT Press, 1999.

Mekala et al. 2018
Sreenivas Mekala – Vishnu Vardhan Bulusu – Raghunadha Reddy T, *A Survey on Authorship Attribution Approaches*, «International Journal of Computational Engineering Research», 8 (9) (2018), pp. 48-55.

Mikolov – Yih – Zweig 2013
Tomas Mikolov – Wen-tau Yih – Geoffrey Zweig, *Linguistic Regularities in Continuous Space Word Representations*, «Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies», (2013), pp. 746-751. aclanthology.org

Pennington – Socher – Manning 2014
Jeffrey Pennington – Richard Socher – Christopher D. Manning, *GloVe: Global Vectors for Word Representation*, «Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing», (2014), pp. 1532-1543, doi: 10.3115/v1/D14-1162.

Roy 2025
Swastik Roy, *Breaking Language into Tokens: How Transformers Process Information?*, «Hugging Face Blog» (blog), 3 maggio 2025, URL: https://huggingface.co/blog/royswastik/transformer-tokenization-vocabulary-creation

Sanders 2022
Ted Sanders, *How to Count Tokens with tiktoken*, «OpenAI Cookbook» (blog), 16 dicembre 2022, URL: https://cookbook.openai.com/examples/how_to_count_tokens_with_tiktoken cookbook.openai.com

Sennrich et al. 2016
Rico Sennrich – Barry Haddow – Alexandra Birch, *Neural Machine Translation of Rare Words with Subword Units*, «Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics», 1 (2016), pp. 1715-1725, doi: 10.18653/v1/P16-1162.

Silva et al. 2024
Kanishka Silva – Ingo Frommholz – Burcu Can – Fred Blain – Raheem Sarwar – Laura Ugolini, *Forged-GAN-BERT: Authorship Attribution for LLM-Generated Forged Novels*, «Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop», St. Julian's: Association for Computational Linguistics, 2024, pp. 325-337, doi: 10.18653/v1/2024.eacl-srw.26.

Turney – Pantel 2010
Peter D. Turney e Patrick Pantel, *From Frequency to Meaning: Vector Space Models of Semantics*, «Journal of Artificial Intelligence Research», 37 (2010), pp. 141-188, doi: 10.1613/jair.2934

Vaswani et al. 2017
Ashish Vaswani – Noam Shazeer – Niki Parmar – Jakob Uszkoreit – Llion Jones – Aidan N. Gomez – Łukasz Kaiser – Illia Polosukhin, *Attention Is All You Need*, «Advances in Neural Information Processing Systems», 30 (2017), pp. 5998-6008

Yang – Carpuat 2025
Xinchen Yang e Marine Carpuat, *Steering Large Language Models with Register Analysis for Arbitrary Style Transfer*, «arXiv preprint arXiv:2505.00679» (2025). arxiv.org

## Appendice

Riscrittura della *'Citta ragnatela'* in diversi stili.

*Dante*: "D'Ottavia, Cité Sospesa, Verace Rima".

Se fede al mio parlar voi porger vogliate, or dirolla qual s'apre, Ottavia, la cittade, che qual ragna d'ogni trama s'è tessute. Duo monti scoscesi, di pietra e rupe forte, un baratro profondo nel mezzo han fesso; e la città è 'l suo vuoto, in mirabil sorte, a quelle creste ligata, qual nodo e intresso, con funi e catene e ponti che vi ponno andare.
Per traversine di legno, ove l'uom si move, convien andar con passo accorto e lesto, ché 'l piè non scivol tra gl'intervalli, dove s'apre 'l nïente, per sì gran disdetto; o pur con man tenace, a le maglie di canapa, ci si distenda, che 'l cader sia evitato. E sotto, nulla v'è per centinaia e tappa, se non le nubi che 'l vento ha dischiarato; e più giù l'occhio scorge, con vista scema, il fondo del burrone, per lunga e trista via.
Questa è la base sua, rete che sì si frema, e serve a passo e a sostegno, in ogni stia. Ma tutto 'l resto, che non s'eleva in suso, sospeso sta di sotto, con ogne suo peso: scale di corda, amache per l'uom riposo, case fatte a sacco, con pensier non fuso, attaccapanni, terrazzi qual navicelle, otri d'acqua, becchi di fiamma per lume, girarrosti, cesti appesi a sottili fardelle, montacarichi, docce che bagna e consume, trapezi e anelli per li giochi al cielo aperti, teleferiche, lampadari che fan luce e decoro, vasi con piante che 'l verde lor han versi.
Sospesa sull'abisso, la vita lor, sì cheto, men incerta è che 'n altre città, lo giuro: ché san ben essi che la rete, più di quel detto, non può reggere, nel tempo che è futuro.

*James Joice*: "Of Octavia, Spider-Spun, a Gossamer City, a Truth Told".

If ye would but lend ear, ye would, and listen, eh? Now, I'll tell you how it hangs, Octavia, this cittade, a web of every thread, woven, unwoven. Two craggy mountains, stone and stark rock, cleft deep a chasm, a gaping maw, and the city itself, a hollow, a wonder of circumstance, bound fast to those crests, knotted, interwoven, with ropes and chains and gangways to traverse.
Along timbers, sleepers of wood, where man-creature stirs, must one go, footfall light and quick, lest the foot slip between the gaps, where nothingness yawns wide, by grievous mischance; or else, grip tight, hand firm, to the hempen mesh, clinging, lest the fall be unfallen. And below, nothing, for hundreds and hundreds of metres, no solid ground, just clouds scudding by, grey puffs, a glimpse below, further down, of the gorge's floor, a long, sad road.
This, then, its base, a net that shivers, yes, it shivers, and serves for passage and for stay, in every dwelling. But all the rest, what doesn't rise, no, it hangs, hangs down below, with all its sundry weight: rope ladders, hammocks for repose, sack-houses, thought unpoured, clothes-hooks, terraces like small ships, water-skins, gas-jets for glim, rotisseries, baskets dangling by flimsy twine, hoists, showers, oh, the washing, rings and trapezes for games flung to sky, cableways, chandeliers for splendour, pots with plants, green leaves weeping.
Suspended there, over the abyss, their life, these dwellers of Octavia, less uncertain, I swear it, than in other cities, no, not so uncertain. For they know full well, deep in their marrow, that the net, beyond what's said, cannot hold, in the time to come. Ah, the net.

*Alla Romana, stile '610':* "Ottavia: 'A Città Che Sta Appesa, E Fattene 'Na Ragione".

Ahò, se me volete crede, bone. Sennò sticazzi, tanto io ve spiego lo stesso com'è 'sta Ottavia, 'a città-ragnatela. Nun è che mo ve mettete a pensa' a 'na ragnatela de quelle che fa er ragnetto ar comodino, eh? Questa è roba seria, fatta a modo suo.
Allora, immaginateve du' montagne, ma proprio du' bestie, che manco l'alpinisti co' le corde d'oro ce s'arrampicano. E in mezzo? Un burrone che te fa venì 'a vertigine solo a guarda' er disegno. E la città? Ma che, sta sopra? Sta sotto? No, sta dentro er buco, proprio in mezzo ar niente, tutta attaccata a 'ste du' creste co' 'na matassa de funi, catene e passerelle, come se fosse 'na bancarella ar mercato, ma appesa all'aria.
A camminacce poi, è 'na goduria! Su 'ste traversine de legno, devi sta' 'na cifra attento a nun mette er piede in mezzo, sennò fai un tuffo che manco a Cinecittà. Oppure te devi aggrappa' alle maglie de canapa come 'na scimmia impaurita, che è pure divertente, se nun c'hai da fa' altro.
E sotto? Sotto c'è er paradiso de chi nun c'ha voglia de fa' gnente: centinaia de metri de vuoto, co' qualche nuvoletta che passa e basta. Ogni tanto, se c'hai 'na vista da falco, s'intravede er fondo der burrone, ma è più che altro per dì: "Aò, c'è un fondo". Insomma, la base de 'sta città è 'na rete. 'Na rete che serve pe' camminacce sopra e pe' nun casca' giù, semplice semplice.
E tutto er resto? Ma che, s'inalza? No, che ve frega? Tutto er resto sta appeso de sotto! Scale de corda, amache pe' fa' 'na pennica co' la morte sotto, case fatte a sacco (per chi ama l'essenziale), attaccapanni (pe' appende i guai), terrazzi che sembrano astronavi (ma nun partono), otri d'acqua (pe' quanno c'hai arsura da vuoto), becchi der gas (pe' nun sta' ar buio), girarrosti (che nun ho capito che ce mettono su), cesti appesi (pe' la spesa che arriva dar cielo), montacarichi (pe' chi se stanca de camminà), docce (che l'acqua poi, bo?), trapezi e anelli (pe' 'n po' de ginnastica acrobatica), teleferiche (pe' i pigri), lampadari (che illuminano er niente), e pure vasi co' le piante che c'hanno le foglie a penzoloni, perché pure loro se so' rassegnate.
E insomma, appesa lì sull'abisso, la vita de 'sti d'Ottavia è meno incerta de quella de l'altre città, ve lo giuro. Perché loro lo sanno. Sanno che più de tanto 'a rete nun regge. E questa, regà, è 'na certezza che te fa sta' sereno, mica balle.

*Agatha Christie in italiano*: "Ottavia: Un Enigma Sospeso sull'Abisso".

Se la mia narrazione desterà in voi una pur minima curiosità, sarò lieto di svelarvi i particolari della città di Ottavia, un luogo che, con ragione, è stato battezzato "la città-ragnatela".
Ci si trova dinanzi a un precipizio, una fenditura profonda che squarcia il fianco di due montagne scoscese, quasi come se la terra stessa avesse ceduto a un'improvvisa, violenta frattura. Ebbene, Ottavia non giace *sulle* creste, né si annida *nel* burrone. Essa *è* il vuoto stesso che si estende tra quelle vette, un vuoto insolitamente occupato, legato alle due sponde rocciose da un'intricata trama di funi, catene robuste e passerelle provvisorie.
Il percorso attraverso questa singolare metropoli impone una cautela non comune. Si procede su traversine di legno, e l'attenzione deve essere costante: un piede malfermo, un attimo di distrazione, ed ecco che lo spazio vuoto si spalanca sotto di sé. La presa alle maglie di canapa diventa un atto riflesso, una necessità vitale per mantenere l'equilibrio.
Al di sotto, per centinaia e centinaia di metri, non vi è altro che aria. Talvolta, qualche nuvola solitaria scivola silenziosamente, offrendo un fugace scorcio sul fondo del burrone, un abisso che

sembra negare ogni solida realtà. È questa la singolare fondamenta della città: una rete, non di pietra o di cemento, ma di fili e collegamenti, che funge al contempo da via di transito e da precario sostegno.

E l'edificazione non segue la logica comune di elevarsi verso l'alto. Al contrario, ogni cosa, ogni elemento della vita quotidiana, è sospeso al di sotto di questa rete primaria. Scale di corda che ondeggiano, amache che dondolano in un sonno precario, case fatte a sacco che paiono contenere segreti inconfessabili, attaccapanni che sembrano custodire non abiti ma indizi. E ancora: terrazzi simili a navicelle pronte a salpare verso l'ignoto, otri d'acqua, becchi del gas che promettono luce nel buio, girarrosti silenziosi, cesti appesi a spaghi sottili, montacarichi, docce che sembrano condurre a nulla, trapezi e anelli per giochi che sfidano la gravità, teleferiche che attraversano il nulla, lampadari che illuminano il vuoto, e vasi con piante dal fogliame pendulo, quasi a voler raggiungere l'abisso.

Sospesa su tale vertigine, la vita degli abitanti d'Ottavia, paradossalmente, si rivela meno incerta che in molte altre città. Essi detengono una verità ineludibile: sanno che, oltre una certa soglia, quella stessa rete, pur così ingegnosa, non potrà più reggere. E in quella consapevolezza, risiede una peculiare, quasi inquietante, forma di certezza.