



Aeneas: un nuovo orizzonte per la ricerca storica

Presentiamo oggi Aeneas, il primo modello di intelligenza artificiale (IA) progettato per la contestualizzazione di iscrizioni antiche. Il suo scopo è aiutare le studioso e gli studiosi di storia antica a interpretare, attribuire e restaurare frammenti testuali con maggiore precisione.

Il team di Aeneas

La scrittura permeava ogni aspetto del mondo romano, dalle imponenti opere imperiali agli oggetti più comuni. Graffiti politici, poesie d'amore, epitaffi, ma anche transazioni commerciali, inviti di compleanno e formule magiche: le iscrizioni offrono agli storici moderni una prospettiva inestimabile sulla diversità della vita quotidiana nell'antica Roma. Tuttavia, questi testi sono spesso frammentari, logorati dal tempo o intenzionalmente danneggiati. Restaurarli, datarli e collocarli geograficamente risulta quasi impossibile senza un contesto adeguato, specialmente quando si confrontano iscrizioni simili.

È con entusiasmo che annunciamo la pubblicazione su *Nature* del nostro [articolo](#) che presenta [Aeneas](#).

Questo innovativo modello di intelligenza artificiale è il primo nel suo genere specificamente realizzato per la contestualizzazione di iscrizioni antiche.

Tradizionalmente, le studioso/i che lavorano con iscrizioni antiche si affidano alla propria esperienza e a risorse specializzate per identificare "paralleli" – ovvero testi che presentano somiglianze in termini di formulazione, sintassi, formule standardizzate o provenienza. Questo processo, sebbene fondamentale, è notoriamente complesso e dispendioso in termini di tempo. Aeneas, invece, accelera drasticamente tale attività: è in grado di analizzare migliaia di iscrizioni latine, recuperando in pochi secondi confronti e paralleli testuali e contestuali. Questo permette agli storici non solo di interpretare, ma anche di approfondire i risultati ottenuti dal modello.

Le potenzialità di Aeneas non si limitano al latino: il modello è adattabile ad altre lingue antiche, scritture e supporti, dai papiri alle monete. Questa versatilità espande la sua capacità di stabilire connessioni attraverso una gamma più ampia di evidenze storiche. Il progetto è stato sviluppato in collaborazione tra Google DeepMind e l'Università di Nottingham, con il contributo di ricercatori e ricercatrici delle Università di Warwick, Oxford e dell'Athens University of Economics and Business (AUEB). Questo lavoro si inserisce in uno sforzo più ampio volto a esplorare come l'IA generativa possa supportare gli studiosi nell'identificazione e interpretazione di paralleli su vasta scala.

Per massimizzare l'impatto di questa ricerca, abbiamo reso disponibile una versione interattiva di Aeneas, accessibile gratuitamente a ricercatori, studenti, educatori, professionisti dei musei e a tutte e tutti gli interessati su predictingthepast.com. A ulteriore supporto della ricerca, abbiamo inoltre reso open-source [il nostro codice e dataset](#).

Le funzionalità avanzate di Aeneas

Prendendo il nome dall'eroe errante della mitologia greco-romana, Aeneas rappresenta un'evoluzione di [Ithaca](#), il nostro precedente lavoro che impiegava l'IA per restaurare, datare e localizzare iscrizioni greche antiche. Aeneas compie un passo ulteriore, supportando gli storici nell'interpretazione e contestualizzazione dei testi, conferendo significato a frammenti isolati, facilitando l'estrazione di conclusioni più ricche e contribuendo a una migliore comprensione della storia antica.

Tra le funzionalità avanzate del nostro modello spiccano:

- **Ricerca di paralleli:** Aeneas è in grado di cercare paralleli all'interno di una vasta collezione di iscrizioni latine. Trasformando ogni testo in una sorta di "impronta storica", il modello identifica connessioni profonde che assistono gli storici nel collocare le iscrizioni nel loro contesto storico più ampio.
- **Elaborazione multimodale:** Aeneas è il primo modello a determinare la provenienza geografica di un testo utilizzando input multimodali. Questo significa che analizza sia il contenuto testuale che le informazioni visive, come le immagini dell'iscrizione.

- **Restauro di lacune di lunghezza sconosciuta:** Per la prima volta, Aeneas può restaurare lacune in testi anche quando la lunghezza del testo mancante è ignota. Questa capacità lo rende uno strumento estremamente versatile per gli storici che si confrontano con materiale gravemente danneggiato.
- **Prestazioni all'avanguardia:** Aeneas stabilisce un nuovo punto di riferimento di eccellenza nel restauro di testi danneggiati e nella previsione della loro datazione e provenienza geografica.



Animazione di un diploma militare bronzeo restaurato dalla Sardegna 113/14 d.C. (CIL XVI, 60).

Come funziona Aeneas

Aeneas è una rete neurale generativa multimodale che elabora testo e immagine dell'iscrizione come input. Per l'addestramento, abbiamo curato un vasto e affidabile set di dati, attingendo a decenni di lavoro di storici che hanno creato collezioni digitali fondamentali, in particolare l'[Epigraphic Database Roma](#) (EDR), l'[Epigraphic Database Heidelberg](#) (EDH) e l'[Epigraphic Database Claus Slaby](#) (EDCS-ETL). Abbiamo quindi pulito, armonizzato e collegato questi record in un unico set di dati leggibile dalla macchina, denominato [Latin Epigraphic Dataset](#) (LED), che comprende oltre 176.000 iscrizioni latine provenienti da tutto l'antico mondo romano.

Il nostro modello impiega un decoder basato su trasformatore per elaborare l'input testuale dell'iscrizione. Reti specializzate si occupano del restauro dei caratteri e della datazione utilizzando il testo, mentre per l'attribuzione geografica vengono utilizzate anche le immagini delle iscrizioni. Il decoder recupera iscrizioni simili dal LED,

classificandole per rilevanza. Per ogni iscrizione, il meccanismo di contestualizzazione di Aeneas individua un elenco di paralleli attraverso la tecnica degli "embeddings": questa tecnica codifica le informazioni testuali e contestuali di ogni iscrizione in una sorta di "impronta storica" che racchiude dettagli sul contenuto del testo, sulla sua lingua, sulla sua origine temporale e geografica, e sulle sue relazioni con altre iscrizioni.

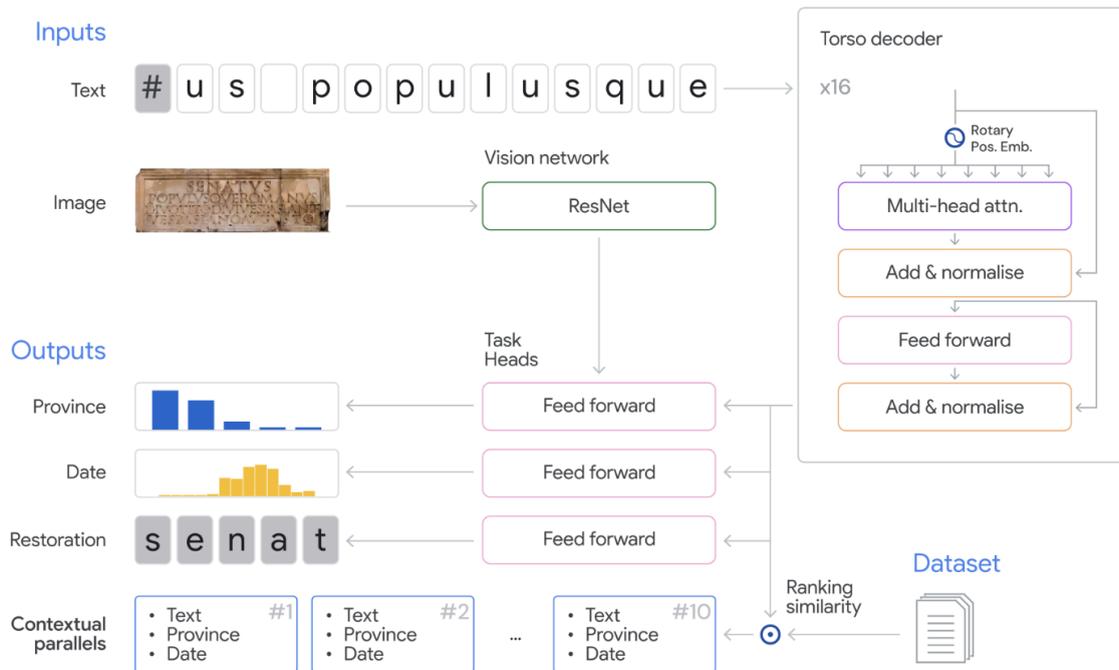
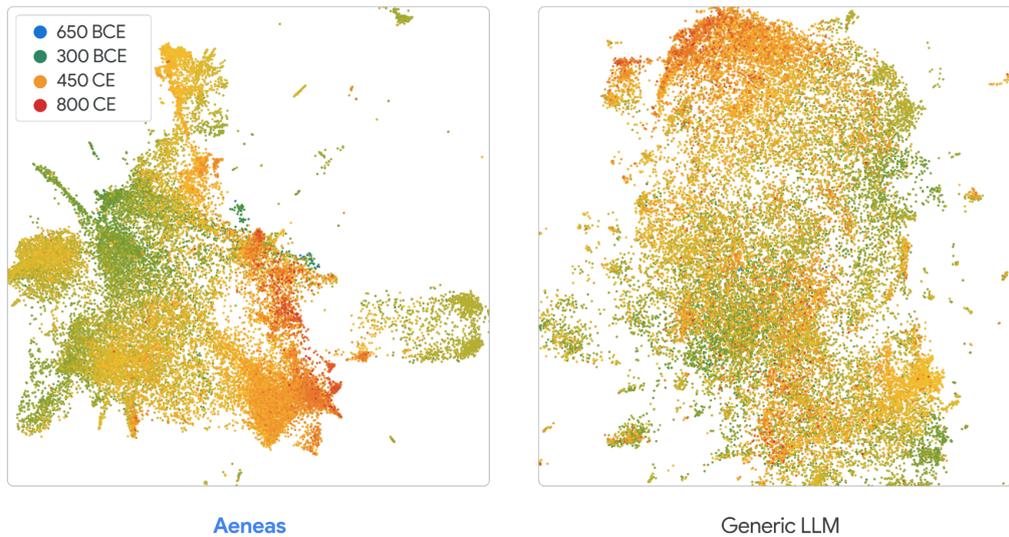


Diagramma dell'architettura di Aeneas che mostra come il modello accetti input di testo e immagine per generare previsioni di provincia, data e restauro.

Prestazioni all'avanguardia

Aeneas raggruppa le iscrizioni in base alla data di scrittura con una chiarezza nettamente superiore rispetto ad altri modelli generici addestrati sul latino, come illustrato nella visualizzazione seguente.

Chronological attribution



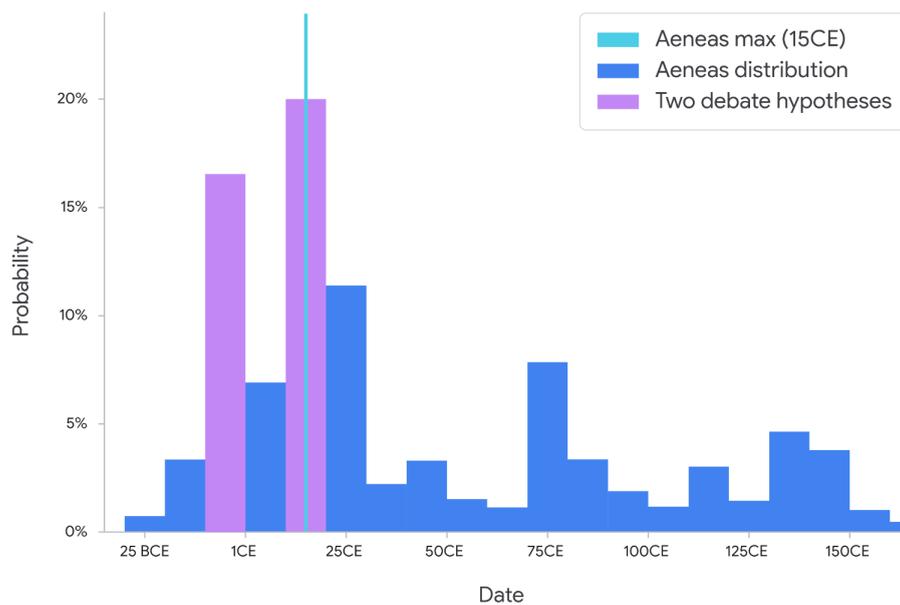
Visualizzazione UMAP (Uniform Manifold Approximation and Projection) che illustra l'attribuzione cronologica degli embedding storicamente ricchi di Aeneas rispetto agli embedding testuali di modelli linguistici generici di grandi dimensioni.

Aeneas è in grado di restaurare iscrizioni danneggiate con una precisione Top-20 del 73% in lacune fino a dieci caratteri. Questa percentuale diminuisce fino al solo 58% quando la lunghezza del testo mancante è sconosciuta, un compito che di per sé è estremamente impegnativo. Il modello mostra inoltre il proprio ragionamento in modo interpretabile, fornendo mappe di salienza che evidenziano le parti degli input che hanno maggiormente influenzato le sue previsioni. Grazie all'utilizzo di dati visivi, Aeneas può attribuire un'iscrizione a una delle 62 province romane antiche con una precisione del 72%. Per la datazione, il modello colloca un testo entro 13 anni dagli intervalli di date forniti dagli storici.

Una nuova prospettiva sui dibattiti storici

Per valutare le capacità di Aeneas nell'ambito di un dibattito di ricerca attuale, abbiamo sottoposto al modello una delle più celebri iscrizioni romane: le *Res Gestae Divi Augusti*, il resoconto in prima persona dell'Imperatore Augusto sui suoi successi.

La datazione di questa iscrizione è stata a lungo oggetto di discussione tra gli storici. Invece di prevedere una specifica data fissa, Aeneas ha prodotto una distribuzione dettagliata di date possibili, evidenziando due picchi distinti: uno minore intorno al 10-1 a.C. e uno più ampio e definito tra il 10-20 d.C. Questi risultati hanno catturato in modo quantitativo entrambe le ipotesi di datazione prevalenti.



Istogramma che mostra la previsione di attribuzione cronologica di Aeneas per le Res Gestae, che modella i dibattiti accademici sulla datazione di questa famosa iscrizione.

Aeneas ha basato le sue previsioni su sottili caratteristiche linguistiche e marcatori storici, come titoli ufficiali e monumenti menzionati nel testo. Trasformando la questione della datazione in una stima probabilistica fondata su dati linguistici e contestuali, il nostro modello offre un approccio nuovo e quantitativo per affrontare dibattiti storici di lunga data. Ancora più significativamente, Aeneas ha recuperato numerosi paralleli rilevanti da testi legali imperiali legati all'eredità di Augusto, sottolineando come l'ideologia dell'impero si riproducesse attraverso l'intero mondo romano.

Avanzare la ricerca storica in modo collaborativo

Per valutare l'impatto di Aeneas come strumento di supporto alla ricerca, abbiamo condotto un ampio studio collaborativo tra storiche/i e intelligenza artificiale. Abbiamo invitato ventitré storiche/i che lavorano regolarmente con le iscrizioni a restaurare, datare e localizzare un set di testi utilizzando Aeneas. La nostra valutazione, riassunta nella tabella seguente, dimostra che i risultati più efficaci sono stati raggiunti quando le studiose/i hanno utilizzato le informazioni contestuali fornite da Aeneas insieme alle sue previsioni per il restauro e l'attribuzione di iscrizioni romane.

Method	Restoration 10-char.	Province accuracy		Date	Historian confidence ↑	Potential as research starting point ↑	N° parallels added ↑
	CER ↓	Top-1 ↑	Top-3 ↑	Distance in Years ↓			
Onomastics baseline		13.07%	23.5%	30.4			
Historian	39.0%	27.0%	42.0%	31.3	49.5%		
With Aeneas parallels	33.9%	36.7%	56.7%	21.1	61.1%	75.0%	1.48
With Aeneas parallels and predictions	21.4%	68.3%	78.3%	14.1	70.0%	90.0%	1.58
Aeneas	23.1%	66.7%	73.3%	12.8			

Tabella che mostra le prestazioni degli storici su tre compiti epigrafici (restauro, attribuzione geografica, datazione) utilizzando 60 iscrizioni dal nostro set di test del database. I compiti sono stati eseguiti prima in modo indipendente, poi con le informazioni sui paralleli di Aeneas, o con paralleli e previsioni insieme.

Aeneas ha supportato le storiche/i che hanno preso parte al nostro studio nell'identificazione di nuovi paralleli e ha aumentato la loro fiducia nell'affrontare compiti epigrafici complessi. Le storiche/i hanno evidenziato all'unisono il valore di Aeneas nell'accelerare il loro lavoro e nell'ampliare la gamma delle iscrizioni parallele e confronti più rilevanti.

“

Aeneas' parallels completely changed my perception of the inscription. It noticed details that made all the difference for restoring and chronologically attributing the text.

— Studios* anonim* dalla nostra valutazione

Condividere gli strumenti, plasmare il futuro

Aeneas è progettato per integrarsi armoniosamente nel lavoro di ricerca storica. Combinando le competenze degli studiosi con l'IA, il modello inaugura un processo collaborativo, offrendo suggerimenti interpretabili che costituiscono preziosi punti di partenza per l'indagine storica.

Stiamo anche aggiornando [Ithaca](#), il nostro modello per il greco antico. Da ora in poi, sarà alimentato da Aeneas e includerà la funzione di contestualizzazione, il restauro di lacune di lunghezza sconosciuta e prestazioni complessive migliorate. Abbiamo inoltre co-progettato un nuovo [programma di studi](#) per connettere competenze tecniche e pensiero storico nell'ambiente didattico. Questo programma è allineato con diverse iniziative di alfabetizzazione sull'IA, tra cui il [Digital Competences Framework for Citizens](#) (DigComp 2.2) della Commissione Europea, l'AI [Competency Framework for Students](#) dell'UNESCO e l'anteprima dell'European Commission and the Organization for Economic Cooperation and Development (OECD) [AILit Framework](#).

Il team di Aeneas prosegue la sua collaborazione con diversi esperti del settore, utilizzando Aeneas per far luce sulla storia antica, con ulteriori sviluppi in programma.

Scopri di più su Aeneas

- [Leggi il nostro articolo](#)
- [Prova Aeneas](#)
- [Ottieni il codice e il set di dati](#)
- Leggi il nostro blog in [inglese](#)

Riconoscimenti

La ricerca è stata co-diretta da Yannis Assael e Thea Sommerschild.

Hanno contribuito: Alison Cooley, Brendan Shillingford, John Pavlopoulos, Priyanka Suresh, Bailey Herms, Jonathan Prag, Alex Mullen e Shakir Mohamed.

L'interfaccia web di Aeneas è stata sviluppata da Justin Grayston, Benjamin Maynard e Nicholas Dietrich, ed è alimentata da Google Cloud.

Il programma di studi è stato sviluppato da Robbe Wulgaert, Sint-Lievenscollege, Ghent, Belgio.