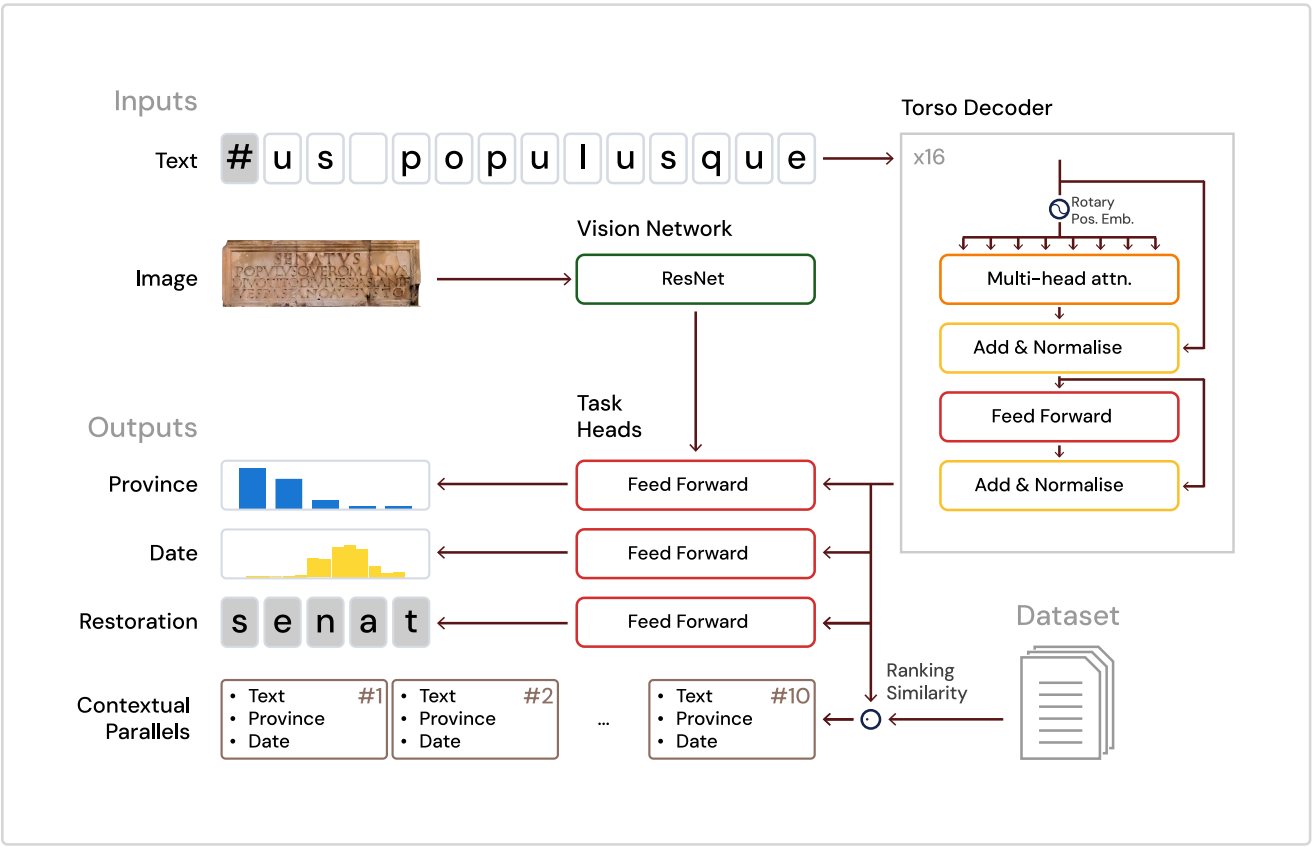
Back to home

AENEAS

Contextualising ancient texts with generative neural networks

Salve! Welcome to Aeneas' interactive interface. Please follow the instructions below to begin contextualising, restoring and attributing Latin inscriptions. You will also find more information on the Aeneas project, links to the article and examples of Aeneas in action.



Google DeepMind · Durham University · University of Nottingham · University of Oxford · Warwick

Human history is born in writing. Inscriptions, among the earliest written records, provide direct insights into the thought, language, and history of past civilizations. Historians capture these insights by identifying textual and contextual parallels—inscriptions with shared phrasing, function, or cultural context—to situate them within broader historical frameworks. This process enables key tasks such as restoring damaged text and determining its geographical or chronological attribution. We introduce Aeneas, the first generative neural network designed to contextualise ancient texts, while also incorporating visual inputs, supporting arbitrary-length text restoration, and enhancing key historical analysis tasks.

Restored bronze military diploma from Sardinia, issued by the Emperor Trajan to a sailor on a warship. 113/14 CE (*CIL* XVI, 60, The Metropolitan Museum of Art, Public Domain).

To evaluate its impact, we conducted the largest Historian-AI study to date, with 23 historians considering Aeneas' retrieved parallels useful research starting points in 90% of cases, improving their confidence in key tasks by 44%. Restoration and geographical attribution tasks yielded superior results when historians were paired with Aeneas, outperforming both humans and AI alone. For dating, Aeneas achieved a 13-year distance from ground-truth ranges.

Method	Rest. 10-ch	Province		Date	Confidenc
	CER ↓	Top-1 ↑	Top-3 ↑	Years ↓	
Onomastics		13.07%	23.5%	30.4	
Historian	39.0%	27.0%	42.0%	31.3	49.5%
W/ Aeneas Context	33.9%	36.7%	56.7%	21.1	61.1%
W/ Aeneas Context & Pred	21.4%	68.3%	78.3%	14.1	70.0%
Aeneas	23.1%	66.7%	73.3%	12.8	

Historians' performance on three epigraphic tasks (restoration, geographical attribution, dating) using 60 inscriptions from the LED test set. Tasks were performed first independently, then assisted by Aeneas' parallels ('w/ Aeneas Context'), or by its parallels and predictions ('w/ Aeneas Context & Pred.'). Metrics include restoration (CER; lower = better), geographical attribution (Top-1/3 accuracy), dating (distance in years), historian's confidence, use of Aeneas' parallels as research starting points, and the number of parallels used.

Method	Rest. Length Fixed	Rest. 20-char		Prov
		CER ↓	Top-20 ↑	Top-1
Onomastics				17.9%
Ithaca	✓	43.5%	44.2%	61.3%
Aeneas	✓	40.5%	46.5%	72.3%
Aeneas	✗	66.1%	32.7%	"

Comparative performance of Aeneas and baseline models on the LED test set across three key epigraphic tasks. Unlike the previous state-of-the-art, which operates under fixed restoration lengths, Aeneas demonstrates flexibility by handling both fixed and unknown restoration lengths.

Aeneas shows how the integration of Science and Humanities can produce transformative tools that assist historians, advancing our understanding of the past.

Aeneas was conceived and researched by Yannis Assael, Thea Sommerschield*, Alison Cooley, Brendan Shillingford, John Pavlopoulos, Priyanka Suresh, Bailey Herms, Robbe Wulgaert, Jonathan Prag, Alex Mullen and Shakir Mohamed. This web experience was developed and built on Google Cloud by Justin Grayston, Benjamin Maynard, and Nicholas Dietrich.*

Read the paper

View on GitHub

Aeneas in the Classroom

Use Aeneas for your research

Enter your Latin epigraphic text, including spaces, in the box below to obtain epigraphic parallels, restore missing characters, and attribute the inscription to its original place and time of writing.

- Use a **question mark** (?) for each character you want the model to predict. Each query can predict up to 20 question marks (consecutive or not).
- Use a **single hash** (#) to predict text sequences of unknown length:
    - You can also adjust the maximum expected length of the restoration using the bar sliders.
    - The total number of characters predicted across both unknown (#) and known (?) length gaps cannot exceed the maximum length you set.
    - Your text should not have any consecutive unknown length restoration characters (e.g. ##), or an unknown length restoration character adjacent to a fixed length character (e.g. ?#).
- Use a **dash** (-) for any missing sections or characters in your text that do not need restoring. You can also use dashes to pad a short text.

- When performing restoration, you can adjust the sampling temperature, which controls how creative or conservative the restoration outputs are. Lower is more like a formulaic funerary text, higher is more like a dedicatory verse inscription.
- There is no need to resolve Latin abbreviations. To restore longer sequences, input inscription images, or inspect more restoration hypotheses, please refer to the Colaboratory notebook.

???ovio tabaliaeno luggoni arganticaeni haec mon# possuerunt

60/760

Restoration Sampling Temperature:

1.0

Restoration Max Length:

20

Contextualise and Attribute

Contextualise, Restore and Attribute

Show me an example

Parallels

Show more

To include as much historical context as possible, we extend the list of retrieved parallels to incorporate the validation and test sets. These sets are not used for training and do not

affect the model's predictions.

Dulovio Tabaliaeno Luggoni Arganticaeni haec mon possierunt

Record Number: HD017521
Province: Hispania citerior

Vianeglo Segei ex gente Abilicorum Tiogilus Caesari posit

Record Number: EDCS-05502087
Province: Hispania citerior

L Couticivi L Coutioso Longonaroso Malgeinus Leuri f Arbuens...

Record Number: HD003296
Province: Lusitania
Date: 1–100 CE

L Coutio Narsi f Caenanicus et Icatl

Record Number: EDCS-34300069
Province: Lusitania

Soncano Caelenio Alloni Occel

Record Number: EDCS-08200069
Province: Hispania citerior
Date: 1–100 CE

Restoration outputs

Restoration hypotheses

Aeneas' Top-20 restoration hypotheses ranked by probability.

Input text:	???ovio tabaliaeno luggoni arganticaeni haec mon# possuerunt
1.	entovio tabaliaeno luggoni arganticaeni haec monumentum de suo possuerunt
2.	entovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
3.	entovio tabaliaeno luggoni arganticaeni haec monimentum de suo possuerunt
4.	entovio tabaliaeno luggoni arganticaeni haec monimentum sibi et suis possuerunt
5.	allovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
6.	licovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
7.	aenovio tabaliaeno luggoni arganticaeni haec monimentum sibi et suis possuerunt
8.	durovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt

Input text: ???ovio tabaliaeno luggoni arganticaeni haec mon# possuerunt

9.	turovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
10.	antovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
11.	allovio tabaliaeno luggoni arganticaeni haec monimentum sibi et suis possuerunt
12.	t bovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
13.	tabovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
14.	licovio tabaliaeno luggoni arganticaeni haec monimentum sibi et suis possuerunt
15.	aenovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
16.	durovio tabaliaeno luggoni arganticaeni haec monimentum sibi et suis possuerunt
17.	ugiovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt
18.	turovio tabaliaeno luggoni arganticaeni haec monimentum sibi et suis possuerunt
19.	antovio tabaliaeno luggoni arganticaeni haec monimentum sibi et suis possuerunt
20.	mogovio tabaliaeno luggoni arganticaeni haec monumentum sibi et suis possuerunt

Restoration saliency map

Saliency maps for each character predicted in Aeneas' top restoration hypothesis, computer in the section above, highlighting in purple shading which unique input text features contributed most to the prediction.

Step	Character	Hypothesis text
1.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_- possuerunt
2.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_-- possuerunt
3.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_--- possuerunt
4.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_---- possuerunt
5.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_----- possuerunt
6.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_------ possuerunt
7.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_------- possuerunt
8.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_-------- possuerunt
9.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_--------- possuerunt
10.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_---------- possuerunt
11.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_----------- possuerunt

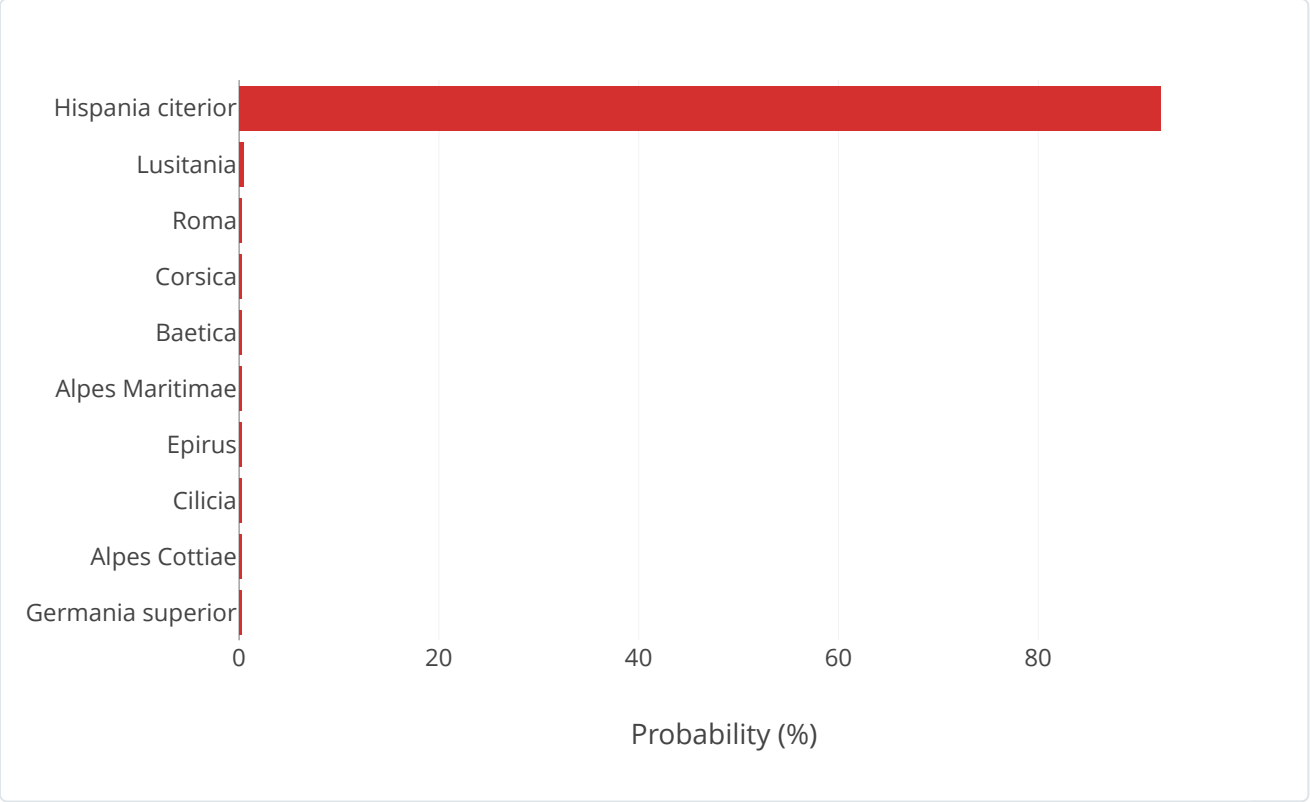
Step	Character	Hypothesis text
12.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_---------- possuerunt
13.	"n"	---ovio tabaliaeno luggoni arganticaeni haec mon_-n---------- possuerunt
14.	"e"	---ovio tabaliaeno luggoni arganticaeni haec mon_en---------- possuerunt
15.	"u"	---ovio tabaliaeno luggoni arganticaeni haec mon_en-u-------- possuerunt
16.	"m"	---ovio tabaliaeno luggoni arganticaeni haec mon_en-um------- possuerunt
17.	"t"	---ovio tabaliaeno luggoni arganticaeni haec mon_entum------- possuerunt
18.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon_-entum------- possuerunt
19.	"-"	---ovio tabaliaeno luggoni arganticaeni haec mon--entum------- possuerunt
20.	"m"	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum------- possuerunt
21.	"u"	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum-----u- possuerunt
22.	" "	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum--- -u- possuerunt
23.	" "	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum -- -u- possuerunt

Step	Character	Hypothesis text
24.	"e"	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum -e -u- possuerunt
25.	"s"	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum -e su- possuerunt
26.	"o"	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum -e suo possuerunt
27.	"d"	---ovio tabaliaeno luggoni arganticaeni haec mon-mentum de suo possuerunt
28.	"n"	-n-ovio tabaliaeno luggoni arganticaeni haec mon-mentum de suo possuerunt
29.	"t"	-ntovio tabaliaeno luggoni arganticaeni haec mon-mentum de suo possuerunt
30.	"e"	entovio tabaliaeno luggoni arganticaeni haec mon-mentum de suo possuerunt
31.	"u"	entovio tabaliaeno luggoni arganticaeni haec monumentum de suo possuerunt

Attribution outputs

Geographical attribution hypotheses

Bar chart and map distribution for Aeneas' Top-10 geographical attribution hypotheses, ranked by probability among 62 Roman provinces. The circle size on the map is directly proportional to the prediction's probability.



Map tiles by Stamen Design (unmodified), under CC BY 3.0. Data by OpenStreetMap, under CC BY SA.

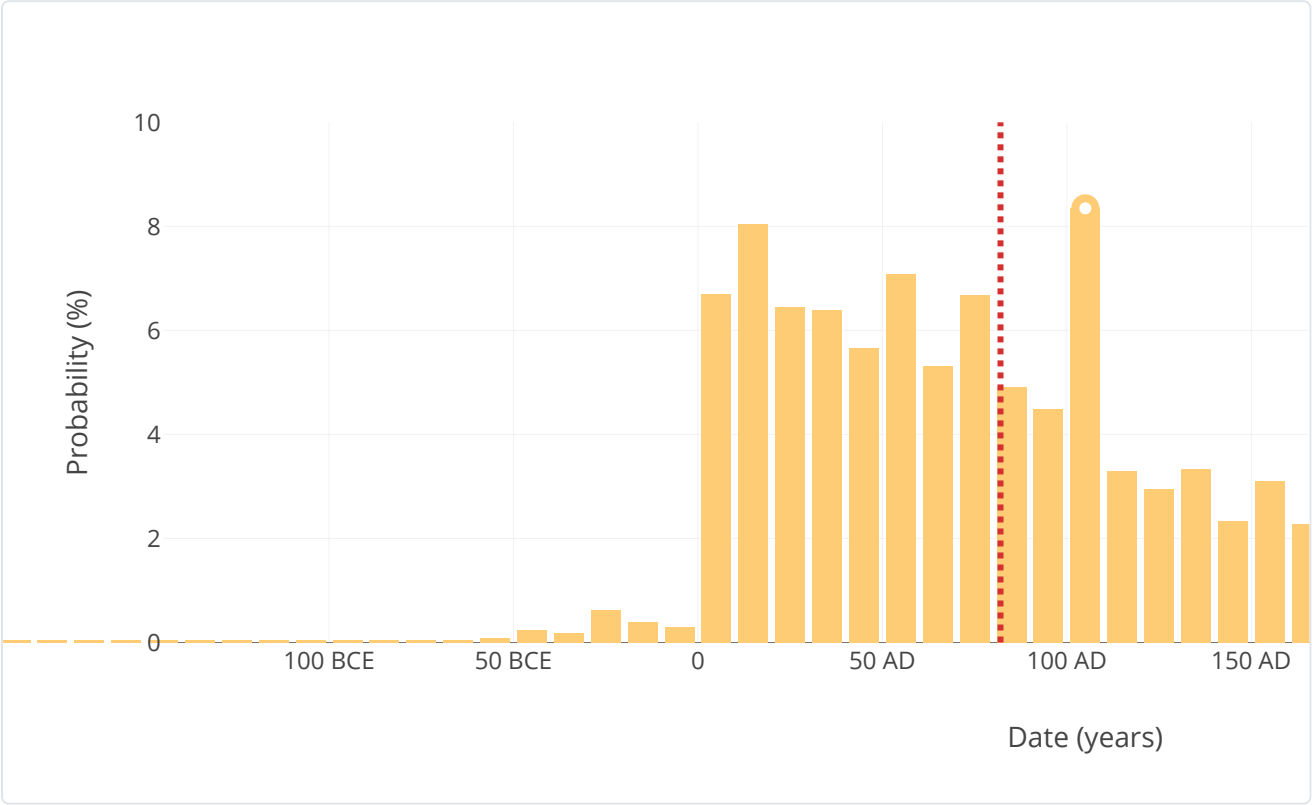
Geographical attribution saliency map

Saliency map highlighting in purple shading which unique input text features contributed the most to Aeneas' top chronological attribution hypothesis, computed in the section above.

???ovio tabaliaeno luggoni arganticaeni haec mon# possuerunt

Chronological attribution hypotheses

Aeneas' chronological attribution hypotheses, visualised as a categorical distribution over decades, in yellow, between 800 BCE and 800 CE. The average of the distribution is depicted with a red line. This enables the handling of date intervals more effectively and aids the interpretability of the hypotheses.



Chronological attribution saliency map

Saliency map highlighting in purple shading which unique input text features contributed the most to Aeneas's top chronological attribution hypothesis, computed in the section above.

???ovio tabaliaeno luggoni arganticaeni haec mon# possuerunt

Responsible AI for the Ancient World: collaborating, not replacing.