

Yes but.. Can ChatGPT Identify Entities in Historical Documents? or, the Postmodern Prometheus

*Oui mais... ChatGPT peut-il identifier des entités dans des documents historiques ?
où le Prométhée postmoderne

Carlos-Emiliano González-Gallardo, Emanuela Boros, Nancy Girdhar,
Ahmed Hamdi, José G. Moreno & Antoine Doucet

Université de la Rochelle, France



TALN/hOUPSh '23 - 5 June 2023

A surrealist painting in a dark, textured, and somewhat desaturated style. The central figure is a mechanical or robotic Prometheus, depicted with a large, dark, metallic body. Its head is a complex assembly of gears, pipes, and a small, human-like face with a beard. The figure is seated at a wooden table, leaning forward and reading a large, open newspaper. Its right arm is extended, holding the paper, while its left arm rests on the table. The background is a wall made of rough, layered stone or wood, with a small, dark, indistinct figure visible in the distance. The overall mood is somber and contemplative.

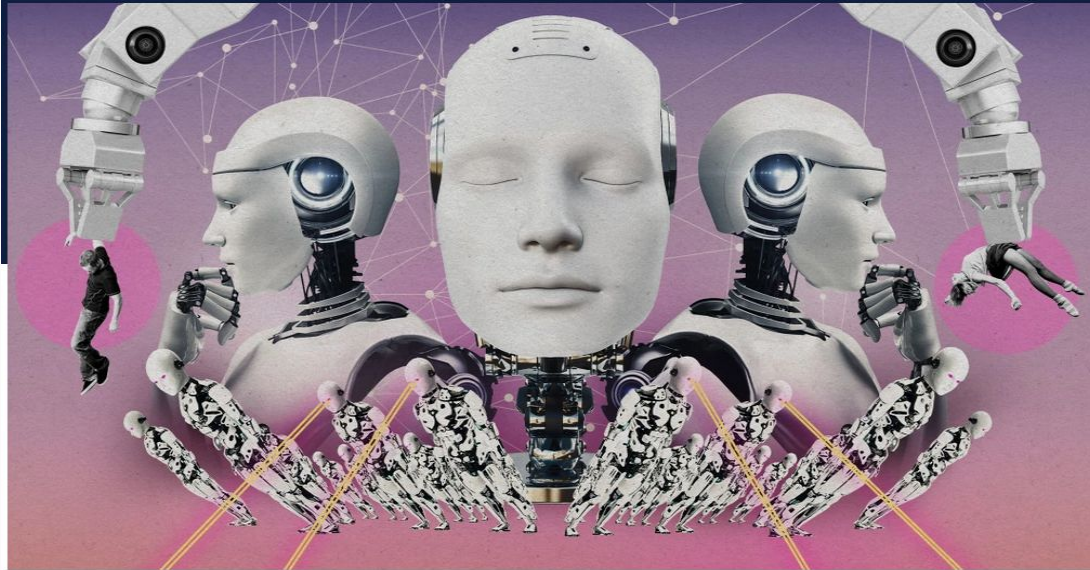
... the Postmodern Prometheus

How to talk about AI (even if you don't know much about AI)

Plus: Catching bad content in the age of AI.

By Melissa Heikkilä

May 30, 2023



STEPHANIE ARNETT/MITTR | GETTY

1. Don't worry about sounding dumb
2. Be specific about what kind of AI you're talking about
3. Keep it real
4. **Adjust your expectations**
5. **Don't anthropomorphize**
6. It's all about power
7. Please, for the love of God, no robots

GPT Assistant training pipeline



A surreal, sepia-toned illustration. A central figure stands, wearing a dark, textured garment. Their head is replaced by a mask composed of several historical faces, including a man with a beard and a woman's profile. The figure's right arm is extended, holding a large, unrolled scroll. Their left arm is bent, holding an open newspaper. The background is a textured, layered wall of various historical documents and maps. The overall tone is historical and academic.

Yes but.. Can ChatGPT Identify Entities in Historical Documents?

Yes but.. Can ChatGPT Identify Entities in Historical Documents?

Carlos-Emiliano
González-Gallardo
carlos.gonzalez_gallardo@univ-lr.fr
University of La Rochelle, L3i
La Rochelle, France

Emanuela Boros*
emanuela.boros@epfl.ch
EPFL, Digital Humanities Laboratory
Lausanne, Switzerland

Nancy Girdhar
nancy.girdhar@univ-lr.fr
University of La Rochelle, L3i
La Rochelle, France

Ahmed Hamdi
ahmed.hamdi@univ-lr.fr
University of La Rochelle, L3i
La Rochelle, France

Jose G. Moreno
jose.moreno@irit.fr
University of Toulouse, IRIT
Toulouse, France

Antoine Doucet
antoine.doucet@univ-lr.fr
University of La Rochelle, L3i
La Rochelle, France

ABSTRACT

Large language models (LLMs) have been leveraged for several years now, obtaining state-of-the-art performance in recognizing entities from modern documents. For the last few months, the conversational agent ChatGPT has "prompted" a lot of interest in the scientific community and public due to its capacity of generating plausible-sounding answers. In this paper, we explore this ability by probing it in the named entity recognition and classification (NERC) task in primary sources (e.g., historical newspapers and classical commentaries) in a zero-shot manner and by comparing it with state-of-the-art LM-based systems. Our findings indicate several shortcomings in identifying entities in historical text that range from the consistency of entity annotation guidelines, entity complexity, and code-switching, to the specificity of prompting. Moreover, as expected, the inaccessibility of historical archives to the public (and thus on the Internet) also impacts its performance.

CCS CONCEPTS

• Information systems → Language models; Information extraction; • Computing methodologies → Natural language processing; • Applied computing → Arts and humanities.

KEYWORDS

Named entity recognition and classification, Large language models, Generative pretrained transformer, Historical documents

1 INTRODUCTION

Since OpenAI¹ released ChatGPT at the thirty-sixth conference on neural information processing systems (NeurIPS) in November 2022², its ability to provide human-like and plausible-sounding answers caused the model to become extremely popular beyond the research community, gaining more than 1 million users in less than one week. ChatGPT is a conversational agent based on GPT-3.5³ (generative pretrained transformer), a large language model (LLM) with more than 175 billion parameters [23]. Given its widespread

popularity and accessibility, the question of how this highly mediated model performs on different natural language processing (NLP) tasks arose already in several fields [3, 24].

LLMs have been leveraged for several years now, obtaining state-of-the-art performance in the majority of NLP tasks, by generally being fine-tuned on downstream tasks such as named entity recognition and classification (NERC) and less in zero-shot settings [21]. Thus, for NERC, but also as a general focus, efforts are dedicated to how to effectively transfer knowledge for domain adaptation by developing cross-domain robust systems and exploring zero-shot, or few-shot learning to address domain and annotation consistency and mismatch in cross-domain settings [14, 15]. Simultaneously, in historical documents (e.g., historical newspapers and broadcasts), NERC faces new challenges apart from domain heterogeneity such as input noisiness, language dynamics and lack of resources [12].

Processes such as optical character recognition (OCR) or optical layout recognition (OLR) affect consistently the transcriptions and thus, this propagated noise influences the precision of NERC systems [4, 5, 18, 22, 29, 30]. Even though the latest developments in deep learning by fine-tuning and pre-training historical LMs brought state-of-the-art results in NERC in historical documents [2, 17], time and domain shifts and resource scarcity remain crucial challenges in learning or reusing appropriate knowledge for NERC. Thus, as expected, LLMs and systems in which they are embedded such as ChatGPT were not explicitly trained for information extraction tasks [26] (e.g., named entity recognition and classification, relation extraction), and moreover, as seen in Figure 1, not specifically with a focus in historical documents.

In this short preliminary work, we conduct an exploratory case study to investigate the potential of ChatGPT, which was trained on a massive amount of Internet data (e.g., Common Crawl, WebText2, Wikipedia) [7] and prompt datasets for reinforcement learning from human feedback (RLHF) [23]. Due to this increase in scale in comparison with previous generative language models such as GPT-3 and GPT-2 [7, 16], the behaviour of the model drastically changed, being considerably more able to perform tasks it was not explicitly trained on than previous models [25]. We would expect ChatGPT to be able to detect entities in historical documents to a certain degree, considering the aforementioned challenges, and therefore, we conduct this study by experimenting with zero-shot NERC and comparing its performance against state-of-the-art systems.

¹This work was done while at the University of La Rochelle, in La Rochelle, France.

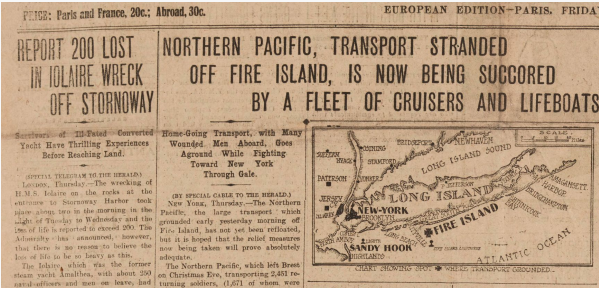
²<https://openai.com/>

³<https://nips.cc/>

⁴GPT-4 was released and integrated into ChatGPT on March 14th, 2023 after the submission of the present work.

A surreal, sepia-toned illustration of a figure with a mask and multiple heads, holding a scroll and a newspaper. The figure has a large, ornate mask with a face and a crown-like structure. It has multiple heads, including a small head on its right shoulder. The figure is holding a large, rolled-up scroll in its left hand and a newspaper in its right hand. The background is a textured, aged wall with various papers and documents pinned to it. The overall style is reminiscent of a historical or artistic painting.

Historical document datasets



ΑΘΗΝΑ.

'ΑΕ! μὲν, ὃ παῖ Ἀαρτίου, δίδορκά σε
 πείραν τιν' ἰχθρῶν ἀρπάσαι θηράμενον
 καὶ νῦν ἐπὶ σκηναῖς σε ναυτικάς ὁρῶ
 Αἴαντος, ἔνθα τάξιν ἰσχύαν ἔχει,
 πάλαι κυνηγετοῦντα καὶ μετρούμενον
 ἰχθυὲ τὰ κεύονα νεοχάραχθ', ὅπως ἰδῆς
 εἴτ' ἔνδον εἴτ' οὐκ ἔνδον, εὐ δέ σ' ἐκφέρει

1. Ἀαρτίου L.A. Ἀαρτίου C¹ Vat. ac. Ἀαρτίου L¹. 6. τὰ 'κείνον L. τῶνδον
 Pal. νεοχάραχθ' L. νεοχάραχθ' C¹ L¹ Vat. ac.

1-3. Athena's eye is ever on Odysseus, and she is now come from Olympus to succour him. Infr. l. 36.
 δὲ μὲν καὶ νῦν] The structure is paratactic; i.e. 'As I have ever seen thee... so now I see thee...' Essay on Language, § 36, p. 66.
 2. (1) 'In quest to snatch some exploit on a foe, i.e. seeking to effect some surprise against a foe.' Or, (2) 'Seeking to foil (or detect) some enemy's attempt.' The latter (2) is simpler, and πείρα is used of the attempt of Ajax, infr. 290, 1057; but the former (1) is on the whole more probable. For Athena does not profess to know the circumstances until l. 36. She asks for information, and only assumes, what is evident, that Odysseus is engaged in some hostile adventure. This aspect of his character appears in the tenth Iliad. Cp. infr. 18, ἰσχυρὸς ἐμ' εἴτ' ἐμὲ δόξα θυγατρὶ] Βάσις καλοῦντ'.
 ἀρπάσαι is to seize, i.e. 'to effect suddenly.' θηράμενον introduces the image of the huntsman continued in l. 5, and combined with that of the hound in ll. 7, 8.
 ἀρπάσαι θηράμενον is substituted for μετροῦμενον, so as to convey the notion of surprise. ἀρπάσαι is an epenegetic infinitive, after which the accusative εἴς is to be resumed. The meaning

of ἀρπάσαι in (2) sup. viz. 'to arrest' is less natural than that given in (1).
 3. σκηναῖς] The σκηνία of the Homeric hero. Cp. infr. 192-3.
 4. ἔνθα. ἔχει] ll. 11, 7, 8; Eur. I. A. 292. This position of Ajax' tent enables him the more easily to steal forth unobserved at last, infr. 690 ff.
 5. κυνηγετοῦντα, which has no object, resumes θηράμενον.
 μετροῦμενον] 'Scanning attentively.' The middle voice marks the mental πείρα is used of the act; not measuring with a line, but scanning with the eye.
 6. νεοχάραχθ'] Ajax has but recently returned, dragging the cattle with him, infr. 296. Odysseus has tracked him so far, but the confused struggle at the tent-door has made it uncertain whether he is not gone forth again.
 7, 8. εἴτ' ἐμ' σ' ἐκφέρει, κ.τ.λ.] Odysseus is like a huntsman who is led to the right point by the scent of a keen Spartan hound. The dog is introduced to complete the image. Cp. infr. 19, 32. Is εἴςινος (1) nom. or (2) gen? εἴςινος occurs in Aesch. Ag. 1093, and the authorities for εἴςινος are late. But the sentence is more balanced if the epithet is taken by hypallage with βάσις, and the abstract noun is somewhat abrupt by itself. Cp. the forms εἴςινος, εἴςινος, εἴςινος.

hipe-2020

pers loc adm.town

par l'honorable cour de justice du Val-de-Travers : le

sieur Daniel Neuron, maître charpentier, bourgeois de Neuchatel.

pers pers.ind comp.name pers pers.ind comp.function pers pers.ind comp.function

ajmc

pers work pers.myth work.primlit scope

Others would render ἀνκρατής, mightysouldered," comparing the description of Ajax in ll. 3. 22 as ἔξοχος Ἀργείων κεφαλὴν ἢ δ' εὐρέας δυόν.

	Tokens	Entities	PERS	LOC	ORG	HumanProd/ object	TIME/ date	SCOPE
NewsEye [19C-20C]	30,458	1,298	463	597	217	21	—	—
*hipe-2020 [19C-20C]	48,854	1,600	502	854	130	61	53	—
ajmc [19C]	5,390	360	139	9	—	80	3	129

Methodology



Prompting ChatGPT

NewsEye

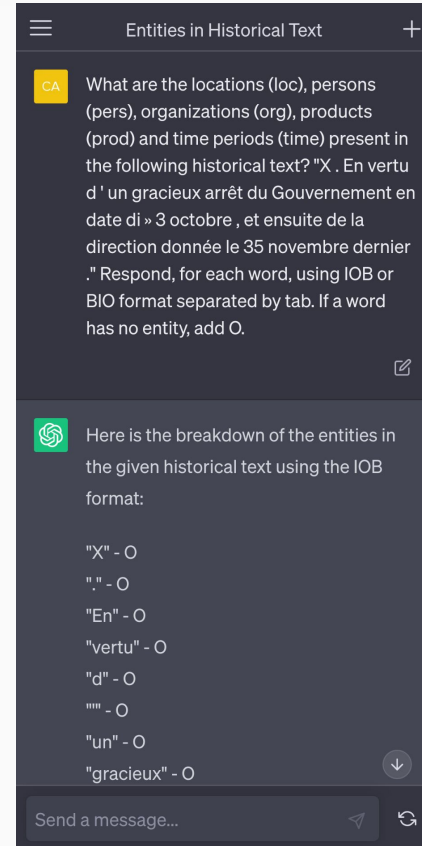
What are the locations (LOC), persons (PER), organizations (ORG) and human productions (HumanProd) present in the following historical text? **{PHRASE}**

hipe-2020

What are the locations (loc), persons (pers), organizations (org), products (prod) and time periods (time) present in the following historical text? **{PHRASE}**

ajmc

What are the locations (loc), persons (pers), time periods (date), human works (work), physical objects (object) and specific portion of works (scope) present in the following historical text? **{PHRASE}**



11th January -
7th February
2023

*Respond, for each word, using IOB or BIO format separated by tab. If a word has no entity, add O. ¹⁰

Some results...

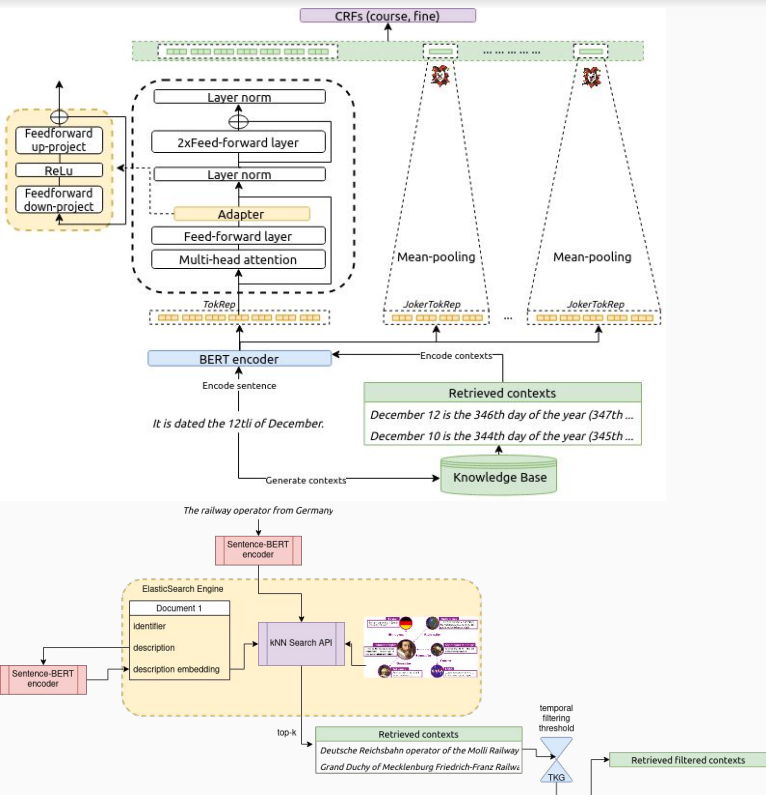


Table 3: Comparative results using the three datasets (micro).

	NewsEye			hipe-2020			ajmc		
	P	R	F1	P	R	F1	P	R	F1
strict									
<i>Stacked NERC</i>	75.0	70.6	72.7	–	–	–	–	–	–
<i>Temporal NERC</i>	–	–	–	76.5	76.5	76.5	84.8	83.9	84.4
ChatGPT	70.9	72.3	71.6	32.5	50.0	39.4	21.8	26.1	23.8
fuzzy									
<i>Stacked NERC</i>	85.4	80.5	82.9	–	–	–	–	–	–
<i>Temporal NERC</i>	–	–	–	86.7	86.7	86.7	90.2	89.2	89.7
ChatGPT	77.8	79.4	78.6	49.0	75.4	59.4	25.5	30.6	27.8

Stacked NERC: Hamdi, Ahmed, et al. "A multilingual dataset for named entity recognition, entity linking and stance detection in historical newspapers.", SIGIR 2021.
Temporal NERC: González-Gallardo, Carlos-Emiliano, et al. "Injecting Temporal-Aware Knowledge in Historical Named Entity Recognition.", ECIR 2023.

A surreal, dark-toned illustration. A figure with a mask that has a single eye and a small, carved face is shown. The figure's limbs are mechanical, with visible gears and joints. They are sitting on a wooden surface, reading a newspaper. The background is a textured, wall-like surface with various patterns and colors. The overall mood is mysterious and industrial.

Limitations

Named Entity Definition

Le sens est donc : « Ai-je parlé trop peu clairement ? » Cf. Antigone, 405 : “Ap’ Eväênda xxi coepñ XéYw ; Eschyle, Agamemnon, 269 : ”H χορῶς λέγω ; [ajmc]

Diagram illustrating Named Entity Definition:

- HumanProd** (purple box) is linked to **SCOPE** (green box) via a dashed arrow.
- SCOPE** is linked to **PERS** (yellow box) via a dashed arrow.

Le sens est donc : « Ai-je parlé trop peu clairement ? » Cf. Antigone, 405 : “Ap’ Eväênda xxi coepñ XéYw ; Eschyle, Agamemnon, 269 : ”H χορῶς λέγω ; [ajmc]

Entity Complexity

┌───┐ ──> PERS

┌───┐ ──> ORG

Beethoven. — Par le Quatuor de la fondation Beethoven : MM.A, Géiosoler violon ; A. Tracol ; 2e violon ; P. Monteux. aito, F. Schnéklud, violon-celle ; César Geloso, pianiste. [NewsEye]

┌───┐ ──> HumanProd

┌───┐ ──> ORG

Beethoven. — Par le Quatuor de la fondation Beethoven : MM.A, Géiosoler violon ; A. Tracol ; 2e violon ; P. Monteux. aito, F. Schnéklud, violon-celle ; César Geloso, pianiste. [NewsEye]

Digitization Errors

└─➔ PERS

« 13659 - 4360.Hxépra. elloug. Ulysse paraît faire allusion à l'amertume des peroles que vient de prononcer À gamemnon ; Agamemnon répond comme si Ulysse avait eu en vue l'amertume de ses propres remontrances, 4866. 'EvOaë' Tlouet, j'en arriverai là, c'est-à-dire, je mourrai, .Dindorf : Kai' αὐτὸς ἴξομαι πρὸς τὸ θάπτειν αὐτόν. » [ajmc]

└─➔ date

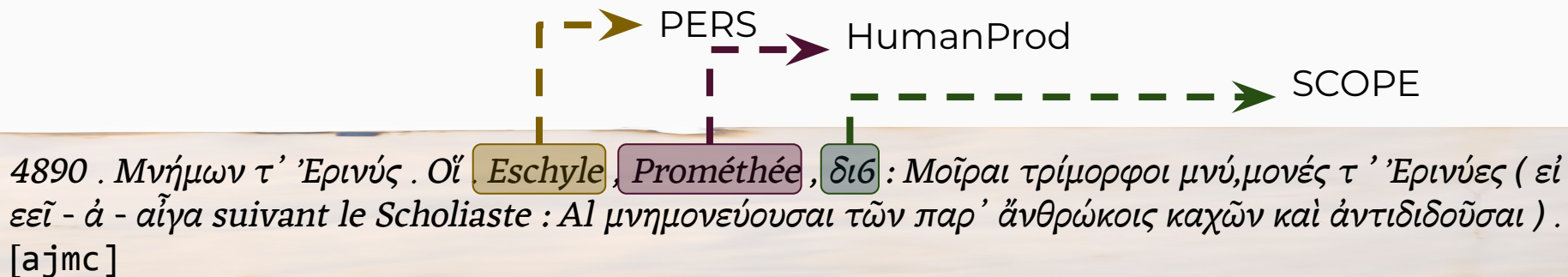
└─➔ object

└─➔ PERS

└─➔ SCOPE

« 13659 - 4360.Hxépra. elloug. Ulysse paraît faire allusion à l'amertume des peroles que vient de prononcer À gamemnon ; Agamemnon répond comme si Ulysse avait eu en vue l'amertume de ses propres remontrances, 4866. 'EvOaë' Tlouet, j'en arriverai là, c'est-à-dire, je mourrai, .Dindorf : Kai' αὐτὸς ἴξομαι πρὸς τὸ θάπτειν αὐτόν. » [ajmc]

Code-switching



I'm sorry, but I'm unable to understand the text you have provided. The text appears to be a mixture of ancient Greek and French, with some references to ancient literature and annotations, which makes it difficult to extract meaningful information. Additionally, the text is not a historical text, but rather a literary text, which also makes it difficult to extract historical entities.



Wrapping up...



Yes but.. Can ChatGPT Identify Entities in Historical Documents?



As an AI language model, I can certainly help you identify entities in historical documents to the best of my abilities. While I don't have direct access to real-time information beyond my knowledge cutoff in September 2021, I can still provide general guidance based on historical data and commonly known entities.



If you provide me with the historical document or relevant text, I can assist you in identifying and extracting entities such as names of people, places, organizations, dates, and other relevant information. However, please note that my responses are based on patterns and information available up until September 2021, so I might not be aware of specific events or developments that have occurred after that time.

Feel free to share the historical document or provide specific information from it, and I'll do my best to assist you in identifying the entities within it.

Merci !

Carlos-Emiliano
González-Gallardo

carlos.gonzalez_gallardo@univ-lr.fr



"A Rembrandt painting of a robotic version of the monster of Frankenstein reading an old newspaper with a background full of knowledge"