

Prompting the Past

Frédéric Clavert (C²DH, University of Luxembourg)

Short paper proposal for DH2024. Available on Github, with code notebook.



Figure 1: “a portrait of Mr Bean as Napoléon Bonaparte” – obviously a more accurate representation of Napoleon than the Ridley Scott movie. Source: lexica.art

Introduction

For several years now, images generated by artificial intelligence (AI)-based systems such as midjourney, DALL·E or StableDiffusion, to give just a few examples, have become commonplace. Like text-generating systems such as GPT, LLAMA or BLOOM, they are based on interaction with users, which notably results in the writing of prompts. The aim of our current research is to examine the various means at our disposal for transforming not only generated images and texts but also prompts into primary sources for historians.

Within a couple of years, the digital history (Salmi (2020)) and digital memory studies (Hoskins (2018)) literature on AI has expanded¹. For instance, Wulf

¹For an extended bibliography on AI and collective memory:
https://www.zotero.org/groups/4874770/ai_collective_memory.

Kansteiner investigated what a specifically trained generative AI for historians could be (Kansteiner (2022)) but without really looking at primary sources produced by these systems. The possibilities and risks of AI, particularly generative AI and especially in Holocaust studies, were analysed in an “open forum” of *Eastern European Holocaust Studies* (Makhortykh (2023)). In particular, the (low) adequacy of the responses of some of the generative AIs to Holocaust research was pointed out (Makhortykh, Vziatysheva, and Sydorova (2023)). Online software can be tested to understand what generative AI systems “know” about history (Hutchinson (n.d.)). Changes – positive or not – in the memorialisation processes have also been analysed (Makhortykh et al. (2023)), rather from the infrastructure angle. Most of those articles are also dealing with ethics, privacy and biases (including biased training dataset). Though research on AI and the past becomes more dense, as far as we know, there is little research on prompts as a source for history or for the study of collective memory.

Questioning the past is the core activity of historians, their basic epistemological operation in the sense that we ask questions to start the process of elaborating new knowledge about the past. The fact that easy to use tools (image or text generative systems) are based on prompts, which are often explicit or implicit questions, should hence get our attention. Those systems are incentives to question our world and the world that was, whereas, as “stochastic parrots” (Bender et al. (2021)), they are a-epistemological: there is no notion of truth, lie or knowledge in the way those systems are working. If we consider what those systems are based on (code, training dataset for instance), what is required to generate an answer (prompts) or outputs (images, texts, etc.), then they are generating numerous primary sources – artefacts that tells us a lot about the societies of the (near) past. We consider here prompts relating to the past as open doors to users’ imagination about the past, but also as the result of a user-machine negotiation (when generative systems are not delivering what the user is expecting, the prompt is modified).

Harvesting data

The first methodological barrier to a research project on prompts related to the past is the making of a corpus. There are several ways we could assemble a database of prompts. A first path would be to use dedicated search engines, such as Lexica. Those search engines poses several problems: either they are basic keywords based search engine and collecting prompts would imply a database of past-related keywords, or they are “semantic” (the case of lexica) and becomes hence black boxes. A further problem with those search engines is that, in our experience, their developers have not kept a crucial metadata: date and time. A second path would be to directly do what those search engines are doing: collecting data on Discord, an instant messaging and VoIP social platform. Discord has indeed become a keypoint for communities to share their uses of large language models or image-generation systems. Companies such as midjourney

have even used Discord as an interface to their system. Data could then be scrapped from discord public servers – after consultation of an ethics review panel² – through a bot to respect Discord terms of use.

The hardest part of the constitution of a corpus is nevertheless to create a balanced corpus – for instance, we can work with data that are not linked to sociodemographic metadata: in other word, we cannot investigate who we are studying – on one side; to define what we mean by "past" on the other side. Many prompts are ambiguous: generated images referring to afrofuturism, for instance, are referring to some future and to the past at the same time.

The lack of sociodemographic metadata could then be supplemented by a complementary qualitative approach, based on interviews with users of generative AI systems. Those interviews could also complete the prompts' data by enlightening us on how users are “negotiating” with the “machine” to get what they expected from it.

Analyzing data

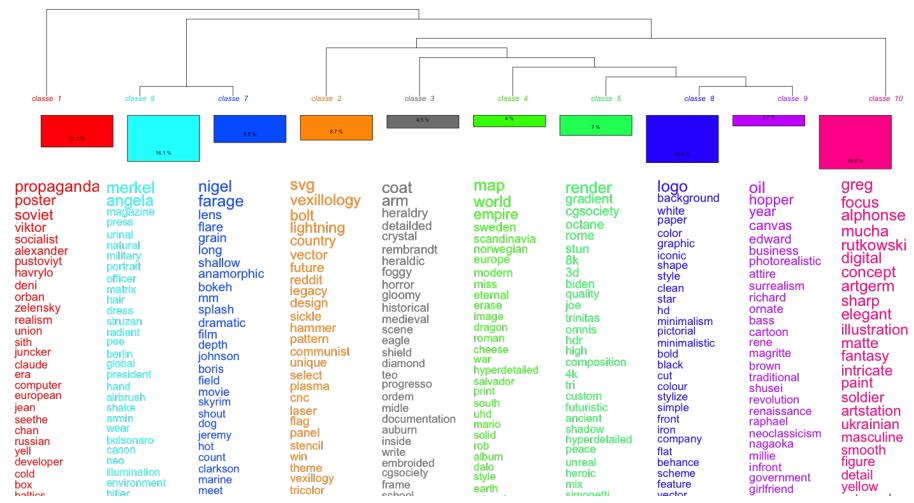


Figure 2: Analysis of the corpus of prompts with IRaMuTeQ

For now, we have collected data from a search engine, `lexica.art`, and will use our small dataset (that can be seen as a test sample) for our distant reading experiments. This small corpus is made of 1908 prompts, collected in March 2023 with the keyword ‘european union’ on `lexica.art`. We chose this keyword to see in which ways users are relating to the European Union’s and Europe’s historical past. The preliminary results show that, whether using `iramuteq`, LDA

²The D/ARC community can offer insights on discord data scrapping that is compatible with the AoIR Ethics Guidelines.

or word2vec³, references to the past are quite important in those prompts. Styles (the style users want for an image) are sometimes ‘historical’ (‘soviet propaganda’ for instance). References to the past are mixed with elements of the current or recent news (‘marine’ for the far-right politician Marine Le Pen in France, ‘nigel farage’ for the Brexit activist, or references to the Russian aggression against Ukraine for instance). Sometimes the notion of ‘Europe’ is linked to a precise period of time, such as the Middle Ages (‘heraldic’). Some historical concepts linked to the European history are also quite visible (‘empire’ for instance). War is very present, but our experience shows that there is a general link that is made between history and war in that kind of corpus.

Conclusion

AI-based generative systems are an occasion for historians to investigate new kind of primary sources, including prompts. Those are presenting some methodological barriers, that can be answered by 1) comprehensive ways to collect data; 2) distant reading of prompts as primary sources. Nevertheless, distant reading of prompts do not give any insights on how users are negotiating with “the machine” when the generated text or image does not fit their (ideological, cultural, etc) expectations. We hence suggest that historians should mix quantitative and qualitative methods, including by doing an oral history of the uses of prompts.

References

- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. FAccT ’21. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>.
- Hoskins, Andrew, ed. 2018. *Digital Memory Studies: Media Past in Transition*. New York: Routledge.
- Hutchinson, Daniel. n.d. “What Do AIs ‘Know’ About History? A Digital History Experiment.” Accessed October 5, 2023.
- Kansteiner, Wulf. 2022. “Digital Doping for Historians: Can History, Memory, and Historical Theory Be Rendered Artificially Intelligent?” *History and Theory* 61 (4): 119–33. <https://doi.org/10.1111/hith.12282>.
- Makhortykh, Mykola. 2023. “Open Forum: Possibilities and Risks of Artificial Intelligence for Holocaust Memory.” *Eastern European Holocaust Studies* 0 (0). <https://doi.org/10.1515/eehs-2023-0053>.
- Makhortykh, Mykola, Victoria Vziatysheva, and Maryna Sydorova. 2023. “Generative AI and Contestation and Instrumentalization of Memory About the

³a notebook is available on github: https://github.com/inactinique/DH2024_prompting_the_past/ – as this is an ongoing research, the code will evolve until the conference.

- Holocaust in Ukraine.” *Eastern European Holocaust Studies*, November. <https://doi.org/10.1515/eehs-2023-0054>.
- Makhortykh, Mykola, Eve M. Zucker, David J. Simon, Daniel Bultmann, and Roberto Ulloa. 2023. “Shall Androids Dream of Genocides? How Generative AI Can Change the Future of Memorialization of Mass Atrocities.” *Discover Artificial Intelligence* 3 (1): 28. <https://doi.org/10.1007/s44163-023-00072-6>.
- Salmi, Hannu. 2020. *What Is Digital History?* 1st edition. Medford: Polity.