Prompting the Past: Co-Creating Historical Insight with AI in the Classroom

By **Krisztina Lajosi, Lily Colsenet, Luca Colombini, Iana Corbeanu, Laia Fernandez, Athina Gianneli, Ruben Rossari, Jessica Wenseritt**

1. Why use genAI for teaching about history?

If higher education wants to keep up with the rapidly changing labour market, it will have to **teach students to experiment and learn with AI tools** while strengthening their creativity and critical thinking skills. For all disciplines, including the Humanities, it is essential to understand what GenAI is, its capabilities and limitations, and how it can contribute value to humanities research. In our European Studies BA, I am responsible for designing a learning pathway that encourages students to use **generative AI ethically and critically** as a tool for learning, and as a way **to reflect on how underlying values and assumptions shape knowledge** itself. Therefore, I included an AI component in our first-year course *Constructing Europe,* and organized several discussions with students about their learning about history with AI.

This article, which I co-authored with some of the students who followed my course, is the result of our class discussions and reflections about **how the use of AI has transformed their approaches to learning and reflecting on European History**.

As a general takeaway, we concluded that reading works written by trained historians is incomparably more valuable than reading AI-generated histories because of their originality, explanatory power, rich details, and accurate referencing, which are epistemic values that are not yet reproduced by GenAI tools. At the same time, **GenAI has the potential to transform future historical research,** opening up possibilities to collect, compute, and parse large datasets and to reveal new perspectives on the past. The future historian requires **solid training in "traditional" source criticism and writing, as well as training in using AI for research and communication.**

The following sections will explain how we utilized GenAI, what we learned about its capabilities for generating texts about history, and how it can be implemented to enhance historical knowledge rather than generalizing and obscuring the details that are often key to understanding history.

2. What GenAi is and what it is not?

GenAI seems to be a genie out of the bottle that can perform complex tasks at lightning speed, which would have taken a human hours of arduous work. However, in fact, **genAI is more of a hybrid between a well-trained parrot and a magician behind a curtain**. It is a system whose workings remain opaque to most users, but the more adeptly we use it, the more it learns from and about us, ultimately providing an output that we desire. The more precise the instructions, the more accurate the AI-generated output will be. The AI is not

(yet) a reasoning machine, but an obedient, patient, and well-trained assistant whose mission is to provide you with the results that you want. **It can help you to clarify what you want, but ultimately, the result will depend on your decision-making and creative skills**.

GenAI excels at replicating, identifying patterns, and predicting text based on the indicated needs of the human agent on a grand scale and at high speed. It has great potential, but if not used with controlled intention, it can become a cliché-spitting machine rather than an innovative and insightful assistant. It truly shines when guided by a creative and informed human mind. When employed thoughtfully and with intent, both the AI and the human can benefit from each other so long as the users give the AI precise instructions and the human user is adept at interpreting the results and deciding whether to accept or reject the AI's output.

Can artificial intelligence help students think more deeply about history? The short answer: it depends. When used as a learning tool together with basic academic skills, AI can foster critical thinking and intellectual curiosity. However, when used as a shortcut for learning, it risks leading to what has been described in the literature about AI usage as "brain rot," the passive absorption of AI-generated information without reflection, which can lead to weaker cognitive skills. In addition, it could result in students learning "botshit," hallucinated, untruthful knowledge, which is an obstacle to learning and thinking.

Students learned that AI is not a historian, but a system trained on vast datasets written by historians. Generative AI is not a superintelligent scholar – at least not yet – but functions more like a super archive, capable of retrieving and rearranging endless files without truly understanding them.

3. How did we use GenAI for learning about European History?

In our first-year course on European history, we aimed to explore the possibilities and limitations of AI in the teaching process. From the teacher's perspective, the goal was to teach students basic skills to use AI ethically and critically for learning and research in history, to enhance research-centered learning and critical thinking, and to experiment with integrating ChatGPT in a way that would best incorporate it into the curriculum. From the students' perspective, the GenAI was supposed to contribute to a more active and thorough engagement with the readings and assignments.

Our goal was not to replace traditional skills but to enhance them with AI. We wanted students to experience how AI could sharpen their awareness of bias, perspective, and narrative construction both in machine-generated and human-written texts. Rather than teaching students to master AI tools, I wanted them to understand how AI shapes research and learning itself.

I designed the assignments so that students were asked to engage with the same historical readings twice: first through conventional means (reading, note-taking, discussion), and then through a dialogue with the university's AI chatbot, guided by a pre-set prompt. They compared the two experiences and reflected on what the AI included, omitted, or distorted.

The results were illuminating about the capabilities and limitations of AI. AI's prose was polished and coherent, but often vague, too generalizing, and was frequently dominated by Western and Eurocentric perspectives. It lacked irony, surprise, and emotional depth, precisely those characteristics that make a human-written text engaging and enjoyable to read. Yet this very absence became an opportunity for learning. Students began to see how narratives are built: what is foregrounded, what is erased, and whose voices are amplified or silenced. They also learned how LLMs are constructed and the best ways to use them for learning, rather than relying on their output mindlessly.

Students discovered that questioning AI is, in itself, a form of critical thinking. As one noted, *"The complexity and nature of an answer depend entirely on the depth of the question."* A vague prompt yields a vague response; a probing, well-contextualized question produces a more meaningful dialogue. This process encouraged students to refine their inquiries, test assumptions, and think historically about technology itself. As a result, AI ceased to be an oracle and became a partner in the intellectual process.

4. What were the observations of the group?

Many students noted that AI's **omissions** spoke volumes. Racism, colonialism, and the experiences of marginalized groups appeared only faintly or not at all (Lily Colsenet and Jessica Wenseritt). The GenAI did not mention minorities like Romani groups, Jews, or European Muslims, and ignored the role of the Byzantine Empire in shaping Europe (Athina Gianneli).

Eastern Europe, the Baltics, and the Balkans, as well as the Scandinavian countries, did not appear in the AI output, thus silencing entire regions or relegating them to the peripheries.

These omissions were not pre-programmed or intentionally programmed into LLMs, but they reveal **biases in digitized data**: histories with more resources will likely be more prominent in the digital space. There is a historical imbalance built into these LLMs: the data AI draws from reflects centuries of unequal documentation and representation. As one student observed, *"The **absences** in AI's answers are proof of which narratives have been documented the most and which have been ignored"* (Laia Fernandez). In addition, we noticed that more details and in-depth knowledge were provided only when specifically prompted and fine-tuned (Lily Colsenet, Jessica Wenseritt, and Laia Fernandez).

The **storytelling** was very poor and superficial, presenting events as historical facts that suggested some kind of inevitability, rather than highlighting what led up to the event and how to understand it in a specific context (Luca Colombini).

Some students mentioned that AI illustrated the dangers of **confirmation bias**. Because it mirrors our patterns of questioning, it can subtly reinforce our existing beliefs (Ruben Rossari). Others reflected on how to counter this tendency: by deactivating AI memory, using primary sources, and, most importantly, developing independent critical thought (Ruben Rossari et al.).

Finally, the question arose: if AI is the future of learning, how will our historical knowledge look? (Iana Corbeanu). If the general public, as well as students, increasingly rely on the Internet for their historical knowledge, and if the Internet is flooded with AI-generated

histories, the algorithms will not prioritize the works of human historians, which would lead to a decline in historical knowledge and to a "dead Internet." There is a shortage of data for LLMs, and AI companies have started using increasingly synthetic, AI-generated data, which makes the AI less reliable. We concluded that we need to invest more in human-generated and curated knowledge, both for the future of building a reliable AI and for building sustainable historical knowledge infrastructures: Knowledge as a Service (KaaS). As Jody Bailey argues, "Rather than scraping static sources, KaaS creates a living, structured ecosystem of contributions from real users (often experts in their fields) who continuously validate and update content. This approach takes inspiration from open-source communities but remains focused on knowledge creation and maintenance rather than code" (Bailey, TechMonitor, 9 July 2025). This insistence on originality and intellectual autonomy echoes a long philosophical tradition. As one student wrote, invoking Immanuel Kant's *Sapere aude!* ("dare to think for yourself"), critical thinking remains the safeguard against the seductions of fluency and authority, whether human or artificial (Ruben Rossari).

Last but not least, reflecting on how to avoid a dead synthetic past on a dead Internet led us to the following question about access to historical knowledge in the future: **how will AI affect our understanding of the past and our historical memory, and how will digital infrastructures impact what is remembered or forgotten in histories**?

The exercise thus revealed not only the biases of the machine but also the limitations of historical knowledge itself in the digital age. AI can both limit and open up endless opportunities for historical knowledge production. It also raised questions about the future: **how will AI affect our understanding of the past and our historical memory, and how will digital infrastructures impact what is remembered or forgotten in histories**.

5. What are our conclusions, and how do we move forward?

The one lesson history teaches is that we can never return to the world before AI. AI will transform education rapidly in the coming decades. It will render many aspects of our educational systems obsolete and highlight other elements that will become even more important. Critical thinking and creativity will undoubtedly be among the most essential skills of the future. Creative critical thinkers who know how to use AI will have a significant advantage in the job market, as against those who have been well-trained to perform tasks that can be outsourced to AI agents. People with knowledge of and a creative mindset for using AI in specific applications will thrive, while those who do not understand the potential and applicability of AI will likely be left behind.

The experiment demonstrated that AI can both help and hinder academic research and critical thinking skills. When used thoughtfully it can enhance the quality and the speed of research. However, when used as "a magic bullet" to do the thinking for the researcher, it does not deliver well.

AI can be a catalyst for deep learning. When guided by thoughtful pedagogy, it encourages reflection on how knowledge is produced, mediated, and reproduced. It helps students not

only to consume information but also to examine it critically, understand its origins, and recognize what it leaves out and why.

Ultimately, the lesson was not about technology, but about thinking. AI can generate text, but it cannot yield insight. That remains the task of human learners: humans must continue to question, doubt, and interpret the world around them.

As one student concluded, *“AI learns from us, but we must continue to be our own writers and judges”* (Iana Corbeanu).