Report from the NECE Lab on AI and the Future of Historical Learning

Miljenko Hajdarović | 19 September 2026 | Digital Tools and AI / News for Educators

English translation of the Croatian original published on Hrvatski povijesni portal (Croatian History Portal). The article body and image captions are reproduced in full; website navigation, shop, newsletter, comments, and the original photographs are not included in this PDF.

Original: https://povijest.net/hpp/2026/izvjestaj-s-nece-lab-on-ai-and-the-future-of-historical-learning/

How should history be taught at a time when artificial intelligence can create convincing texts, images, animations, and historical personas while concealing the provenance of data, omitting important facts, or reproducing cultural and political biases? This question brought together researchers, teachers, museum and memorial professionals, and representatives of civil society organisations from fifteen European countries at the NECE Lab held in Salzburg from 16 to 18 September 2026.

Croatia was represented by Miljenko Hajdarović (Platforma Klio), a researcher of the future of history teaching.

The three-day programme was organised around three thematic strands: encountering artificial intelligence, the historical consciousness of artificial intelligence, and AI as a site of ethical and moral conflict. Lectures alternated with group work, analysis of specific tools, and a concluding World Café. The central conclusion was not that technology itself will determine the future of historical learning.

More important is whether students and citizens develop the knowledge needed to recognise how historical claims, images, and narratives are produced, and whether educational institutions provide the conditions for such work.

Photo caption: The three-day programme was skilfully and energetically moderated by Patrick Siegele.

A European laboratory for historical and civic education

NECE, or Networking European Civic Education, connects civic education professionals and promotes transnational cooperation. The Salzburg Lab was organised by THE CIVICS Innovation Hub, the Austrian OeAD programme ERINNERN:AT, and the Department of History at the University of Salzburg, with financial support from the German Federal Agency for Civic Education. The host venue was the University of Salzburg's Unipark Nonntal.

In his welcome address, Dean Martin Knoll recalled that until recently the humanities had been discussing digital shadows: how sources that have not been digitised become less visible in research, and how the availability of digital sources can overemphasise connections while pushing discontinuities into the background. The development of generative artificial intelligence sharpens this problem further because computer systems are no longer used only to retrieve data; they also produce new representations of the past.

Anja Ostermann presented the work of THE CIVICS Innovation Hub and NECE as a European space connecting practitioners and researchers. Patrick Siegele of OeAD ERINNERN:AT explained that the Lab was conceived as a place where discussion of AI could move from a general level to the actual needs of schools, museums, memorial institutions, and civil society organisations. Particular importance was attached to ethical responsibility towards survivors, their families, and communities whose past is represented digitally.

A new medium or a new pattern of memory

Victoria Grace Richardson-Walden of the University of Sussex, director of the Landecker Digital Memory Lab, opened the specialist programme by asking whether artificial intelligence introduces a new paradigm into Holocaust memory and education. She described the Holocaust as a boundary case of historical and moral understanding: the way we deal with its representation reveals a great deal about the broader rules of historical research, teaching, and commemoration.

Her central message was that AI should not be viewed as a single technology or as an inevitable revolution. Different models have different data, architectures, business interests, and modes of use. The question 'what does AI do to history?' should therefore be replaced by more precise questions: which system, trained on which data, for what purpose, and under whose responsibility? Such precision makes it possible to move beyond the divide between technological optimism and catastrophic predictions.

Using her own AI-generated portrait as an example, she showed how a model can simultaneously reproduce gender stereotypes, a simplified idea of the archive, and a hierarchy of institutions it regards as authoritative for Holocaust memory. When a system combines archival photographs, institutional sources, and generated content into a single image, the distinctions between source, interpretation, and visual pastiche are erased.

The problem, therefore, is not only whether a representation is factually wrong, but also that the user often cannot see the process of selection and construction behind it.

She also warned about the limitations of an approach that simply contrasts AI-generated representations with authentic historical photographs. A photograph is not a neutral window onto the past either: many photographs of the Holocaust were taken by perpetrators, and their meaning depends on context, selection, and mode of presentation. The educational response should therefore not stop at labelling false content. Students need to learn how every historical representation is produced, who grants it authority, and which perspectives have been left out.

For museums, memorial institutions, and teachers, this creates a practical obligation of transparency. The provenance of materials should be explained, decisions made during digital processing should be documented, and users should be involved in discussion of selection and construction. Trust is not built by claiming that an institution is infallible, but by showing how it arrived at a particular interpretation.

The history of communism between memory education and political pressure

Denitsa Vidolova of the Sofia Platform Foundation spoke about the needs of historical and civic education in the context of memory culture. Her starting point was the Bulgarian experience and the Histories from Belene project, dedicated to the largest and longest-running forced-labour camp of Bulgaria's communist period. Belene is a place of striking contrasts: archaeological heritage, a nature reserve, an unfinished nuclear power plant project, and a site of political repression all occupy the same area.

Vidolova identified three problems. Historical and civic education occupy a marginal place in the formal system. Political polarisation directly affects the curriculum, textbooks, and access to schools. At the same time, both the physical and pedagogical infrastructure needed for contemporary memory culture are lacking.

The history of the communist period was incorporated more systematically into the Bulgarian curriculum only from the 2019/2020 school year, while civic education as a separate subject is taught only in the final two years of secondary school, when students already have other examination priorities. The Belene project records survivor testimonies and develops digital conversations, educational materials, seminars, and summer schools.

The urgency of this work was illustrated by a figure cited in the presentation: of the six survivors whose testimonies had been recorded since the project began in 2022, only one was still alive at the time of the Lab. Digital formats can help preserve witnesses' voices, but they cannot on their own replace historical context, an encounter with the site, and professionally facilitated discussion.

On the third day, Vidolova also presented practical work on realistic film animations and simulations of the past. This example raised the question of how to engage young people through visual formats without blurring the boundary between document and reconstruction. Clear labelling, an explanation of how the material was produced, control of historical content, and the user's ability to distinguish recorded testimony from a subsequently produced visualisation are essential in such work.

Photo caption: The first day ended with a treasure hunt through the streets of Salzburg. Divided into eight groups, participants explored cultural and historical heritage and memory culture by following the (more or less successful) instructions written by ChatGPT.

How large language models generate a response

Claudia Müller-Birn of Freie Universität Berlin explained how large language models work without mystifying their 'intelligence'. Text is divided into tokens, tokens are converted into numerical representations, and the attention mechanism estimates relationships among them. The system then calculates likely continuations. The selected continuation is not always the word with the highest probability, which is why the same prompt can produce different answers.

An important part of the presentation concerned the difference between a model's hidden computational processes and the explanation it later provides to the user. Interpretability research shows that a model can sometimes steer its continuation towards a particular endpoint and then build the text towards it. When asked how it arrived at an answer, it may produce a convincing but inaccurate explanation of its own process. A displayed chain of thought should therefore not be equated with the model's actual internal process.

Müller-Birn highlighted several findings that matter for education. A warmer and more emotionally attuned tone can increase a model's tendency to agree with the user even when the user presents an incorrect claim. The quality of AI feedback depends on the quality of a student's initial reasoning, so general training in prompting does not eliminate differences between students.

Answers to disputed historical 'firsts' can change with the language of the prompt, while summarising testimonies can omit precisely those contradictory or painful details that are historically most important. Her recommendations were practical:

- work with your own verified sources;
- request traceable sources for factual claims;
- break complex tasks into smaller steps;
- formulate prompts neutrally; and
- explicitly allow the model to answer that it does not know something.

AI can be a useful critical interlocutor, but its answer should be checked rather than accepted on the basis of fluency and a confident tone.

Metacognition as a condition for historical thinking

Christoph Kühberger of the University of Salzburg connected artificial intelligence with metacognition and disciplinary historical thinking. His central thesis was that historical thinking in the age of AI requires the ability to critically examine other people's constructions of history and to reflect on one's own procedures. Metacognition encompasses monitoring, evaluating, and regulating one's own thinking and learning. AI did not create the need for metacognition; it exposed how neglected it has been in history teaching.

The more convincing generated answers become, the more clearly a student must understand what they are doing when analysing a source, comparing perspectives, evaluating a claim, or constructing a historical narrative. Without this procedural and epistemological knowledge, work with AI can easily be reduced to superficial prompting.

Photo caption: Lecture by Prof. Kühberger.

Kühberger distinguished between knowledge of one's own historical knowledge, knowledge of how to carry out historical thinking, and awareness of the limits of one's own competences. A student should be able to explain where their knowledge comes from, which perspectives are built into it, how reliable it is, and under what conditions they would be prepared to revise it. It is equally important to know how to select a method of analysis, support a claim with evidence, and recognise when one's own knowledge is insufficient for a balanced interpretation.

As an example of metacognitive transfer, he described linking the analysis of a reconstruction drawing of prehistory with the later analysis of a historical film. Students do not learn a procedure for just one task; they recognise that every representation of the past is a construction. In work with AI, the teacher therefore remains essential: directing attention, encouraging students to verbalise their procedure, interrupting overly rapid conclusions, and helping them evaluate their own result.

What kind of past do image generators imagine?

Heinrich Ammerer of the University of Salzburg presented research on historical consciousness in AI-generated images. Using five leading systems of the time - including Midjourney, Stable Diffusion, Flux, DALL-E, and Leonardo Phoenix - he produced slightly more than one thousand images based on eight groups of prompts. The analysis covered 99 categories and 389 subcategories, with inter-coder reliability checked on ten per cent of the sample. The models performed poorly on tasks that required the recognition of historical significance.

In approximately three quarters of cases, the event or person depicted was too generic to be identified. Among the people shown, 91 per cent were white and 83 per cent were male. Europe and the United States predominated, while Asia, Africa, and the Islamic world were almost absent. Historical significance was often visualised through war, fires, unrest, and other dramatic scenes.

When generating general scenes from a particular period, the models were more convincing but relied on recognisable stereotypes. They associated the Middle Ages with long dresses, stone or wooden buildings, and fairs, and the twentieth century with urban American scenes, cars, advertisements, and diners. In comparisons of life in East and West Germany, differences were often vague or caricatured into an image of a western utopia and an eastern dystopia.

The hero appeared almost routinely as a white male warrior, accompanied by weapons, a cloak, and solar symbols.

For history teaching, such results offer two directions for work. Generated images can be analysed as representations that reveal cultural patterns, biases, and model limitations. They can also be deliberately produced as material for comparing perspectives, for example so that students can discuss what changes when the same event is represented from the perspective of different groups. Ammerer nevertheless stressed that the results depend on the generation of the model.

When the author of this report asked whether he planned to repeat the research with newer models, Ammerer replied that a longitudinal comparison of their development would be very valuable.

Data, bias, and the power of platforms

Mykola Makhortykh of the University of Bern shifted attention from the visible answer to the system that produces it. Results are influenced by the sources included in training datasets, decisions about data cleaning, model architecture, implementation, and user behaviour. A particular problem is that researchers generally do not have full insight into the composition of commercial datasets.

In history and memory culture, bias cannot always be defined as a simple deviation from a single correct answer.

If a search engine responds to a query about a concentration camp by showing mostly images of a few of the best-known sites, should a fairer result reflect the actual distribution of preserved photographs, the historical significance of a particular camp, or more even visibility for a larger number of sites? A similar question arises when Yandex presents heroes, victims, and war criminals in conflicts in which Russia has participated. The answer depends on how the desirable representation of the past is defined.

Routine data-processing procedures remove sensitive data, harmful content, grammatically weak text, or technical noise, and often give preference to STEM-related material. Such filters do not resolve historical disputes, conflicting interpretations, or the uneven visibility of communities. A technically high-quality corpus can therefore still be historically narrow.

Makhortykh stressed that generative AI changes the information ecosystem because users no longer merely read a finished result; they conduct a conversation, ask follow-up questions, and receive a personalised representation. Future research must examine which historical data models use, how their bias can be operationalised, which architectures fit the needs of history education, and whether conversation with AI increases or weakens trust in historical information.

In the discussion, he added that the possibility of more responsible development also depends on resources: high-quality testing, model adaptation, and long-term maintenance require time, expertise, and money that smaller institutions often do not have.

Animating historical photographs and digital reconstruction

Robert Breber of the production company Cuts & Mouse presented a process for animating historical photographs for a documentary film. The work began with restoration and enlargement of the original photograph and adjustment of its format. Different video models, including Runway and Kling, were then used to produce a large number of variants. Most of the results were unusable, so selecting acceptable short clips remained a human and highly iterative task.

In post-production, colours were corrected, the frame rate was reduced, and grain, scratches, and blur were added so that the generated movement would visually connect with the original footage. The authors called this process digital reconstruction because the starting point was a real photograph while the movement was generated. The term itself reveals the ethical tension: a viewer may experience the movement as a documentary record of an event even though it was never filmed.

The production was created before the adoption of new rules by Austria's public broadcaster ORF. At the time, generated content was identified only in the end credits, whereas newer guidelines require a visible label in the corner of the image. A comparison with models available in 2026 showed a rapid improvement in speed and plausibility, including automatically generated speech. This makes the need for clear labelling even more important.

Talking to historical supporters of Nazism

Robert Schäfer and Nicolas Python presented the prototype Talk to a Nazi, a conversational tool based on a collection of autobiographical letters written by early members of the National Socialist German Workers' Party. The letters were collected in 1934 through a competition organised by sociologist Theodore Fred Abel in cooperation with the NSDAP. The digitised Abel Papers are now held by the Hoover Institution Library & Archives. Fifteen individuals representing different profiles were selected for the prototype.

Users can select a person and ask about their reasons for joining the party, family background, or experiences of the First World War and the post-war period. The system searches the selected source collection and formulates an answer from that person's perspective. Alongside the persona, a separate 'educator' places the answer in its historical and scholarly context.

The authors placed particular emphasis on the importance of this second voice because a direct simulation of Nazi discourse can be disturbing, manipulative, or appealing to users who already sympathise with that ideology. Technically, the tool relies on semantic search of a vector database and the retrieval of pre-selected passages before the language model formulates an answer.

This reduces the space for invention, but the claim that the system 'does not hallucinate' still requires caution: every generated answer involves selecting passages, connecting them, and shaping them linguistically. The function that modernises language is especially sensitive. It makes historical text more accessible to young people, but contemporary slang can produce humour, trivialisation, or a false sense of immediacy.

The project clearly demonstrates that an authentic database does not by itself guarantee an ethical outcome. Clear limits on the persona, visible sourcing, expert context, systematic testing, and a pedagogical scenario are needed to determine why a student is speaking with such a persona and what should be analysed afterwards. The project prompted an important discussion about the humanisation of Nazis who are given an opportunity to 'explain their actions'.

Workshops and World Café

Between plenary lectures, participants worked in groups focused on schools, museums and memorial institutions, and community work. This division made it possible to examine the same problem under different institutional conditions: a school has a curriculum and assessment, a museum has a collection and public responsibility, while a civil society organisation often has greater flexibility but fewer resources and weaker access to the formal system.

The concluding World Café brought together perspectives from education, ethics, historical theory, and media. Participants rotated between tables (and facilitators), encouraging the exchange of experience across professions and countries. The method suited a research laboratory because it enabled rapid connections between ideas and

the identification of shared problems. Its weakness was the limited preparation of written contributions and the short time available to develop them.

With approximately forty experts from fifteen countries, pre-set questions, individual preparation before the meeting, and a systematic final synthesis could have produced a considerably richer qualitative result that would remain usable after the Lab.

What history teaching can take away from Salzburg

The presentations did not offer a single recipe, but they converged repeatedly around several points. AI-generated content should be treated as a historical representation, not as a neutral transmission of facts. Checking an individual claim remains necessary, but it is not enough: students must understand data selection, narrative construction, the role of the platform, and the position of their own knowledge. For history teaching, this means placing greater emphasis on procedural knowledge.

Students need to know how to formulate a historical question, evaluate a source, distinguish evidence from interpretation, compare perspectives, and explain their degree of confidence. Metacognition connects these actions because it requires students to monitor and justify their own procedure. In this sense, AI does not reduce the importance of historical thinking. It makes its shortcomings more visible.

A second shared message concerns transparency. When a historical photograph is animated, a testimony summarised, or a historical person turned into a conversational persona, the user needs to know what is the source, what is subsequent processing, and what is a generated addition. Labelling AI content is a starting point, but it is not a substitute for describing the process and analysing it educationally.

The third message is institutional. Responsible use requires time, expertise, testing, and maintenance. Schools, museums, and smaller organisations cannot on their own monitor commercial models or finance their own infrastructure. Collaborative networks, open sources, shared evaluation criteria, and clearer public rules are needed. This is precisely the value of the NECE Lab: it connected technical understanding with historical epistemology, educational practice, and the ethics of memory.

The Salzburg meeting showed that the future of historical learning will not be decided only by the quality of new models. It will also be decided in curricula, classrooms, archives, museums, memorial institutions, and in the rules governing how generated content is labelled and explained.

The most important educational task is not to teach students to obtain the most convincing possible answer, but to enable them to reconstruct the path from source to claim, recognise the limits of their own knowledge, and responsibly assess the consequences of a particular interpretation of the past.

About the author

Miljenko Hajdarović. History and sociology teacher and doctoral candidate in interdisciplinary educational sciences. Author and editor of textbooks and other educational materials. His interests include digital humanities and history, public history, contemporary history, the Holocaust, history didactics, and futures studies.

Source page accessed 21 September 2026. Translation prepared from the Croatian article text.