



The Use of Generative Artificial Intelligence in the Digitisation of Printed and Manuscript Documents and Its Contribution to Historical and Archival Education

KLÁRA RYBENSKÁ, SYLVA SKLENÁŘOVÁ

Abstract: Objectives – This study examines how contemporary generative language models can support archival and historical work with Czech handwritten texts, focusing on transcription and basic preliminary analysis, and it outlines key limitations and ethical implications for educational use in archival science and digital humanities.

Methods – A qualitative case study was conducted using seven modern personal handwritten Czech texts from the 1980s and 1990s (a poem written by a child, love poems, a school test, and study notes). Three widely available tools in their free versions (ChatGPT, Claude, and Copilot) were tested using identical task instructions. Outputs were comparatively evaluated with regard to transcription accuracy, content and stylistic interpretation, and recognition of selected formal features of the texts. The empirical comparison was complemented by a critical review of relevant scholarly literature and reflection on authenticity, data integrity, epistemic security, and personal data protection.

Results – Claude achieved the best overall performance, followed by ChatGPT, while Copilot produced substantially weaker results in the tested tasks. Across tools, interpretation and analysis proved more challenging than transcription, and outputs included errors and over-interpretations that require expert verification.

Conclusions – Generative language models can function as supportive tools for transcription, preliminary analysis, and didactic work, but they cannot replace professional archival or historical expertise. Responsible use requires critical human supervision and explicit attention to ethical and data-protection considerations.

Keywords: generative artificial intelligence, large language models (LLMs), handwritten texts, HTR (Handwritten Text Recognition), archival science, historical research, manuscript transcription, AI ethics, digital humanities, AI in education



INTRODUCTION

At present, the fields of archival science and historical research are undergoing significant transformations. A fundamental role in these changes is played by advances in so-called generative artificial intelligence technologies, which also offer new possibilities for historians, archivists, and other professionals working with historical sources (Spina, 2023). When generative artificial intelligence is being discussed, it is understood as artificial intelligence techniques (hereinafter referred to as AI) that are used to create or produce new content, such as texts, code, images, videos, or sounds (Gartner, 2024). Through so-called deep learning, generative AI can recognise meaningful patterns and relationships within large databases, which enables it to generate various forms of content (Mello et al., 2023). One of the key technologies of generative AI is natural language processing, which includes not only the analysis of language but also the understanding of human language (Leviner, 2023).

The case study presented in this article works with personal handwritten texts from the 1980s and 1990s (thematically focused, for example, on school-related and leisure texts), which in our research serve as model data for testing the possibilities of generative AI in the transcription (and, where applicable, basic understanding) of handwritten text in the Czech language. These texts are understood as representative examples of modern handwritten documents with which

historians or archivists may work in the future, and at the same time as didactic material suitable for teaching in the fields of historical studies and archival science. The conclusions of the study are therefore deliberately limited to this type of material and do not apply to all historical texts.

Generative AI can automate the transcription of historical texts, significantly accelerating the digitisation process and increasing the accessibility of documents for researchers, students, and the general public. Another critical technology in the field of generative AI for working with texts is undoubtedly OCR (Optical Character Recognition), which is used to convert scanned text into machine-readable text. This technology is widely applied in digitisation projects to transform physical documents into digital formats that can be edited and searched. For historical texts, OCR technology is beneficial, as it is capable of transcribing even old and damaged documents, and in some instances also handwritten or faded texts (Berchmans & Kumar, 2014; Philips & Tabrizi, 2020; Ferro et al., 2023). Although the problem of OCR for machine-printed text has been more or less resolved (depending on the language and legibility of the given text, the quality of the paper, and similar factors), human handwriting exhibits an almost infinite range of scripts and styles, and its accurate recognition represents a considerably more complex task, which is addressed by HTR (Handwritten Text Recognition) technology, that is, the recognition of handwritten text (Dietrich, 2021).



Generative AI can also be used for transliteration, which is particularly important for texts written in different languages and scripts, as it enables easier comprehension and analysis of such texts. OCR and HTR, therefore, play a crucial role in preserving and making accessible valuable documents. However, the process of transcribing these texts manually can be not only time-consuming but also potentially prone to human error. In this respect, AI tools, similarly to the earliest OCR technologies, have undoubtedly brought about a revolution in the field of text transcription by significantly reducing the time and effort required to transcribe various types of texts, including those of historical value. These advanced technologies thus have the potential to change the way textual sources are analysed, interpreted, and transcribed.

Moreover, AI may contribute to a further revolution not only in historical research but also in education as a result of its ability to analyse and interpret the content of documents. Owing to the exponential development of AI technologies, even more advanced capabilities can be expected, such as context recognition, the extraction of additional information from documents, or the automatic correction of already-transcribed OCR texts (Memon et al., 2020). Finally, it is also necessary to mention that AI facilitates the creation of digital archives and repositories, thereby ensuring the preservation and accessibility of texts for future generations. Digital archives of this kind

can not only serve as resources for scholarly research but may also contribute to the broader dissemination of historical knowledge among the public (Sabharwal, 2015; Spina, 2023).

The integration of AI into text transcription through OCR and HTR holds significant promise for expanding our understanding of the past and preserving cultural heritage. Collaboration between AI specialists and historians, archivists, and other experts will undoubtedly lead to further innovations that will ultimately advance the fields of archival science and history, enriching our knowledge of the past. Moreover, the use of AI tools in text transcription can help overcome language barriers and thus support international cooperation, not only in historical research (Spina, 2023; Ferro et al., 2023), but also in the field of education and pedagogical practice.

The article was produced with the support of the *Philosophical Faculty of the University of Hradec Králové*. Artificial intelligence tools (e.g. Grammarly and DeepL) were used in the preparation of this article for language editing and proofreading of the English version of the text.

1. RESEARCH OBJECTIVES AND MOTIVATION

Within this study, we focus on the possibilities of using contemporary generative artificial intelligence tools for the transcription and analysis of handwritten texts in the Czech language. We deliberately



work with handwritten materials that are not stored in any digital database or public repository and have not been published. By doing so, we aim to minimise the likelihood that these texts were included in the training data of the models used, and thus to approximate a situation in which AI operates on “new” material by an unknown author.

For the case study, we selected seven handwritten texts from the personal archive of one of the authors and her sister. The texts were produced when the authors of the materials were approximately between 11 and 22 years of age, thus capturing the development of handwriting from childhood script through adolescence to early adulthood. The corpus includes both continuous handwritten texts written in cursive script and one text written in block letters. This selection enables the observation of not only individual changes in a person’s handwriting over time, but also a comparison of the handwriting of two individuals at the same age. However, their comparison and development do not constitute the objective of the study. A detailed description of the texts that were selected is provided in the methodological section (see Section 3).

From a palaeographic perspective, this material is of interest in that, within a relatively small sample, it demonstrates changes in letter forms, script size, slant, and overall legibility depending on age, writing proficiency, and the writing material used. Moreover, the author of most of the texts belongs to among those writer with variable handwriting – her script

changes depending on the length of the writing, line spacing, or paper format – which represents an additional layer of complexity for AI tools.

From the perspective of modern diplomatics, we are interested in the extent to which AI tools are capable of identifying basic formal and content-related characteristics of a text – for example, the type of document, its structure, possible purpose, the writing instrument used, or its basic thematic focus. Such analysis may in the future assist in the preliminary classification of undated or poorly described materials within a particular time period or context. Given that the authors are very familiar with the texts, it is possible to assess the AI outputs not only in terms of linguistic accuracy but also in terms of their correspondence with the real circumstances of the documents’ creation.

For the general public, this research has potential significance in that it demonstrates the possibilities (and limitations) of using easily accessible AI models for transcribing family, personal, or professional handwritten materials without the user having to undergo specialised palaeographic training. At the same time, it addresses the question of the extent to which such transcription is reliable and when, by contrast, expert intervention is necessary.

In the future, consideration may also be given to utilising AI in the examination of the authenticity of handwritten documents, for example, in the verification of signatures or the detection of forgeries, which are areas addressed by diplomatics.



However, this dimension already goes beyond the scope of the present research and is indicated only as a possible direction for further investigation.

The aim of the research is therefore to evaluate the current capabilities of selected generative artificial intelligence tools in the reading, transcription, and fundamental content-related and formal analysis of handwritten texts that are not publicly accessible and that represent modern personal handwritten materials.

For this research, we selected three widely accessible tools available in free, commonly used versions: ChatGPT (via its public chat interface), Claude, and Microsoft Copilot. Data collection took place in late 2024 and early 2025; therefore, the study does not cover newer model releases introduced later or paid/enterprise deployments. Nevertheless, it is reasonable to expect that more advanced model variants and enterprise-grade implementations may yield stronger performance – for example newer generations of Claude (e.g., Claude Sonnet) or GPT-4–based tools integrated into Microsoft 365 (Copilot). The study evaluates not only transcription accuracy but also the models' ability to interpret content and style and to identify selected formal features of handwritten texts. It further compares differences across tools and discusses their suitability for archival practice, digital humanities, and education. Particular attention is paid to ethical issues, especially authenticity and data integrity when working with systems capable of generating new or modified content.

2. KEY CONCEPTS AND DEFINITIONS

Generative artificial intelligence, as already indicated in the introduction, refers to a broad range of advanced models and machine learning techniques capable of generating new, original content – such as images, audio, video, text, code, and similar outputs – based on learned patterns. These models are trained on diverse, large-scale datasets in a way that enables them to capture the fundamental structures and characteristics of the input data. This allows them to produce highly realistic and varied outputs that closely imitate the properties of the original data (Goodfellow et al., 2020; Dietrich, 2021).

The potential of AI for archival and historical research is enormous. It promises a revolution in the ways documents and other types of sources are processed, analysed, and interpreted. Through these technologies, archivists and historians can automate and streamline the digitisation and transcription of both printed and handwritten texts. This facilitates preservation and increases the accessibility of sources not only for specialists but also for the wider public. As a result, it can significantly expand the reach and social impact of archival and library collections, making them more accessible for study and public engagement (Arthur et al., 2004). Nevertheless, digitisation – and in particular text recognition using artificial intelligence (AI) – cannot avoid problems related to accuracy and correct transcription. The reading and transcrip-



tion of printed texts is generally less problematic, while handwritten texts pose greater challenges.

One of the primary applications of AI in archival research is the digitisation and processing of documents. Numerous advances in the fields of deep learning and computer vision have led to the development of models that are capable of accurately recognising and transcribing both printed and handwritten text (see below). In this way, they overcome the limitations of traditional optical character recognition (OCR, *optical character recognition*) techniques (Gozalo-Brizuela & Garrido-Merchán, 2023).

HTR (*handwritten text recognition*) technology is generally associated with more challenges than OCR for several reasons, including the following:

- *handwriting variability* – each person writes differently, which makes it challenging to create a universal model for handwritten text recognition (Baheti, 2022; Hedau, 2024);
- *irregularity and complexity of handwritten characters* – unlike standardised printed fonts, handwritten characters may differ significantly in their size, spacing, and alignment. Such irregularities require more sophisticated algorithms for correct text interpretation (Hedau, 2024);
- *handwriting quality* – some handwriting is challenging to read or even completely illegible even for expert researchers, which significantly complicates automatic character recognition (Nockels et al., 2022);
- *linguistic and historical variations* – different languages or dialects may have distinct writing conventions which require specialised models for each language group. Older texts often also contain archaic or otherwise unusual language, which language models may then transcribe or translate inaccurately (Mondal & Jawahar, 2022; Nockels et al., 2022; de Sousa Neto et al., 2024);
- *condition and nature of historical documents* – the physical condition of historical documents can present a significant problem, as they may be damaged, faded, or written in various historical scripts, which significantly complicates their automatic recognition (Nockels et al., 2022; Leviner, 2023);
- *annotations and overlays* – historical texts often contain marginalia, text overlaps, doodles, insertions, and other disruptive elements that significantly reduce the quality of input data for HTR and thus hinder the automatic recognition of handwritten text (Alkendi et al., 2024);
- *technological requirements* – modern HTR systems rely on advanced deep learning models and require large, well-annotated training datasets that represent different handwriting styles and languages. These training processes are computationally demanding and resource-intensive. Finally, it is also essential to consider that the quality of results is susceptible to the quality of digitisation (Baheti, 2022).

As indicated earlier, advances in deep learning and computer vision have led to



the development of models that, in some instances, are capable of accurately recognising and transcribing handwritten text. In this way, they complement – and in some cases, overcome – the limitations of traditional OCR/HTR techniques, which is remarkable because they can be further fine-tuned on specific datasets. Generative AI models can be trained on large datasets of (not only) historical documents, thereby acquiring representations of the fundamental characteristics of texts. Using this knowledge, it is possible to generate synthetic data that faithfully imitates the appearance and properties of real documents. This enables the development of more robust and more universal tools for document processing (Epstein et al., 2023). AI can also be used to increase accessibility and engagement with archives, and, over time, to automate the processing of handwritten and printed texts (Bartz et al., 2020).

However, despite these promising advances, the integration of AI into archival or historical research also introduces additional challenges that need to be carefully considered. In this respect, it is considered crucial to focus on the ethical implications of utilising AI-generated content, particularly in the context of historical narratives, to maintain the authenticity and integrity of historical data (Kadaruddin, 2023).

Both historians and archivists are gradually integrating AI into their workflows, thereby operating within a complex environment of ethical, practical, and technical considerations (Frontoni et al., 2022). One of the main concerns is

undoubtedly the potential misuse or misinterpretation of AI-generated content, which may lead to the dissemination of inaccurate or misleading information. Moreover, generative language models can be trained, for example, to produce persuasive texts that may be indistinguishable from those written by humans. This raises questions about the reliability and trustworthiness of information obtained from these models, particularly when dealing with historical texts and primary sources (Ma et al., 2023). The integration of AI into archival and historical research must therefore be carefully balanced against the need to preserve the authenticity and integrity of historical data. Although digitisation increases accessibility, it must not compromise the original artefact or introduce unintended distortions or misinformation into the output (Colavizza et al., 2021).

One of the more prominent directions in this area is the use of generative adversarial networks (GANs, *Generative Adversarial Networks*) and related models for the synthesis of realistic document images. These models learn the visual characteristics of historical documents, such as paper texture, ink distribution, or other typical features of handwriting. As a result, they can generate synthetic documents that may, at first glance, closely resemble the originals. Such synthetically generated data can subsequently be used to expand training datasets for models intended for document analysis or recognition. An example of a synthetically generated handwritten document is illustrated in Figure 1.

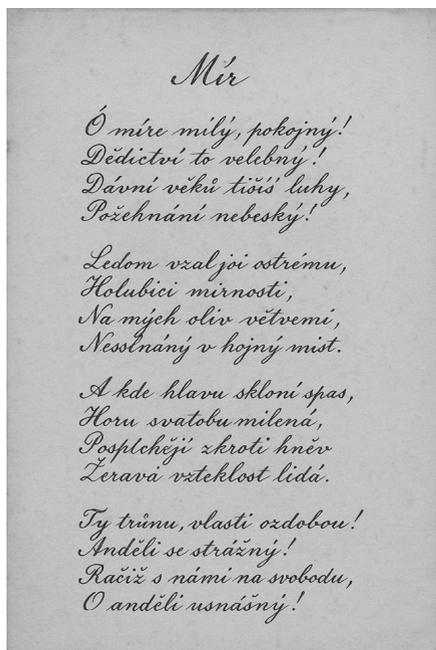
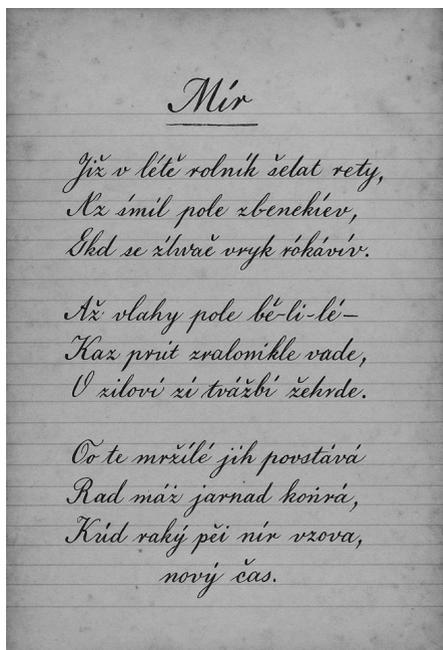


Figure 1 Although this does not represent the most realistic AI-generated output (it is easy to recognise that the text contains typographical errors and that, at present, this still allows us to infer that AI created the document), technologies of this kind have nevertheless progressed significantly.

Synthetic documents of this type can be used to expand training datasets for document analysis and recognition models, thereby improving their performance and robustness. In addition to document digitisation, AI can also be applied to tasks such as transcription, summarisation, and text generation. Language models trained on large corpora of historical and modern texts are capable of producing coherent and contextually relevant transcriptions of handwritten or printed documents, thus

streamlining the conversion of analogue materials into searchable digital formats. This makes the content more accessible to both researchers and the general public. Furthermore, these generative AI models can be used to summarise key information and insights contained in historical documents, providing historians and archivists with valuable analytical tools for navigating and contextualising their collections (Epstein et al., 2023; Gozalo-Brizuela & Garrido-Merchán, 2023).



3. CHARACTERISTICS OF THE HANDWRITTEN TEXTS THAT WERE ANALYSED

A total of seven texts, varying in length and type, from the private archive of one of the authors were used for the research. The corpus includes both continuous texts written in cursive script and a text combining cursive and block letters. In one case, the entire manuscript is written in block letters. This text was deliberately included to ensure that, in comparison with handwritten scripts, the difference in the handling of transcription would be observable.

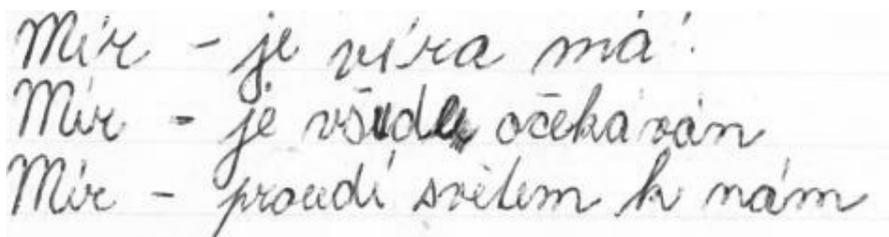
The texts were written at the ages of 11, 17–18, and approximately 22 years. In adulthood, the handwriting already incorporates block letters within the written text. Each handwritten sample, even among those by the same author, exhibits different characteristics, such as variations in script slant, letter

size, and letter shape. This represents varying levels of difficulty for AI tools and may effectively constitute multiple types of handwriting for the models to process.

From the perspective of content and genre, the corpus includes:

- authorial poems written at approximately 11 years of age and around 20 years of age,
- a literature test from secondary school,
- excerpts and notes from books and textbooks from the period of university studies,
- humorous dialogues.

All the texts originate from the second half of the 1980s and the 1990s. The oldest text – a poem – was created in the second half of the 1980s, while the remaining texts date to the period after 1990. For each of the handwritten documents that was analysed, the article provides a brief description and a short sample of handwriting that the AI models were required to transcribe.



Text No. 1 represents a child's poem entitled *Mír (Peace)* written by the author. Given that she attended primary school in the 1980s, its content is slightly influenced by the atmosphere and educational context of that period.



Ska stíka' dolí jo troci'
 Plnu mede mnou pusek koci'
 A ja' smutí o duši máim

Text No. 2 is an untitled poem written at around 20 years of age. It was written in the mid-1990s. In this text, a distinctly different handwriting is already evident, including the non-use of certain forms of cursive script, such as the block form of the lowercase 't' or the uppercase 'S'. It is a love poem in which a young woman comes to terms with the departure of her partner.

1. Kdo je duchovním otcem následujících hrdinů?

Myriam - Klíma ✓

poručík Borůvka - Škvařil ✓

Borek Trojan - Páral ✓

Jiří Křík Kavalír - Hoff ✓

Text No. 3 is part of a test in Czech literature from secondary school. The same author wrote it, although the handwriting differs significantly in both shape and size. The text dates from the mid-1990s. This text is written using a combination of cursive and block letters.

ALEOGRAFIE (z řec. PALAIOS = starý a GRAFFEN = psát)
 // - starý stěra' písmo, psaní mýj písmo jako tehleto
 paleografie 'letníka' = psaní písmem letníka obrubí bez
 ohledu na pautit' psaní
 tím, že slouží a popisuje písmo v čase i prostoru, má spínní čist
 a interpretovat písmo písmem a zeraže mýj písmo psaníto
 kulturní, poli' letníka a hosp. psaní te' letníka' spolínost

Text No. 4 is a set of notes from university studies in auxiliary historical sciences; it is an excerpt from the book *Vademecum* created for study purposes in the mid-1990s.



Daj mi kousek svého ja'
a ja' ti dam zase svůj.
Budem jisté krásný pár,
Tak se tolik nestrachnýj.

Text No. 5 is another love poem, written in the second half of the 1990s. Unlike all the other texts, it is written entirely in block letters. The same author wrote it.

AKTA 4

S: maldři 2x psaly na mi.
M: Delej S. Tady v kdu m' no
mni byl.

Text No. 6 was written by a different person. It was written at approximately 17 years of age, also in the second half of the 1990s. It consists of part of a dialogue intended for an entertaining skit.

ONA: TAKŽE
ON: O.K. Boun so
ONA: Ale kach, pschy p'dem
ON: O.K. 20 p'dem 80 so ala

Text No. 7 was written by the second person. The same unchanged handwriting can be observed here, even though the text was written approximately one year later than the previous one. It again consists of part of a dialogue from an entertaining skit prepared for a public performance.



4. PRACTICAL IMPLEMENTATION OF THE RESEARCH

In this section, we present an experiment in which three different large language models – *ChatGPT (OpenAI)*, *Claude (Anthropic)*, and *Copilot (Microsoft)* – were asked to cooperate in the transcription and analysis of the handwritten texts described in Section 3. In all cases, commonly available free versions of these services were used. The experiment was conducted at the end of 2024 and at the beginning of 2025. The aim was to verify their capabilities in the recognition of handwritten text and in the identification and interpretation of textual content. All the models were given identical instructions, deliberately formulated in natural and unedited language that might be typical of an ordinary student or researcher attempting to consult an artificial intelligence system, provided that they have at least a basic awareness of so-called role-based prompting. In addition to qualitative description, we also carried out an indicative quantification of transcription accuracy for three representative handwritten texts (the poem *Mír* – Text No. 1, the love poem – Text No. 2, and the dialogue – Text No. 6), defined as the proportion of correctly transcribed words relative to the total number of words in the text. The aim was not to construct a fully-fledged statistical mod-

el, but rather to complement the qualitative evaluation with a simple quantitative indicator that allows for a clearer comparison of the performance of the individual tools.

The prompt that was provided to all the tools that were tested was written in the Czech language and was as follows:

“Hello, please respond as an expert in palaeography, diplomatics, epigraphy, and history. You are an expert comparable to, for example, [Real Person 1], [Real Person 2], or [Real Person 3], who is also a sphragistics specialist. I will be sending you some texts in JPG format, and I will ask you to provide a transcription and, where appropriate, any additional information that you consider relevant. Do you understand the task?”¹

This formulation was deliberately kept in a less formal style in order to simulate common uses of AI outside the academic environment and at the same time to test the ability of the models to interpret loosely structured requests, similarly to how AI might be addressed, for example, by a student of archival studies or by a user without deeper technical or specialist training.

The texts were presented to the models in the same order, and in exceptional cases it was necessary to adjust the file size in order to preserve legibility. Subsequently, the results of the transcriptions and analyses were recorded and compared.

¹ The so-called role-playing prompt, role-based prompting (also referred to as expert emulation prompting, among other terms), represents a type of instruction that directs a language model to adopt a specific expert identity (Lee & Palmer, 2025). In our case, the model was instructed to act as a specialist in the fields of palaeography, diplomatics, epigraphy, and history. The prompt was deliberately formulated in a less formal and more natural style in order to reflect real-world uses of artificial intelligence in an educational context – for example, by a student who needs to transcribe their own handwritten notes, or by an early-career researcher using AI to consult difficult-to-read manuscripts. The aim was to test how the model would handle not only the transcription itself, but also a user-friendly prompt formulated outside strictly expert frameworks. (ABBYY, 2023)

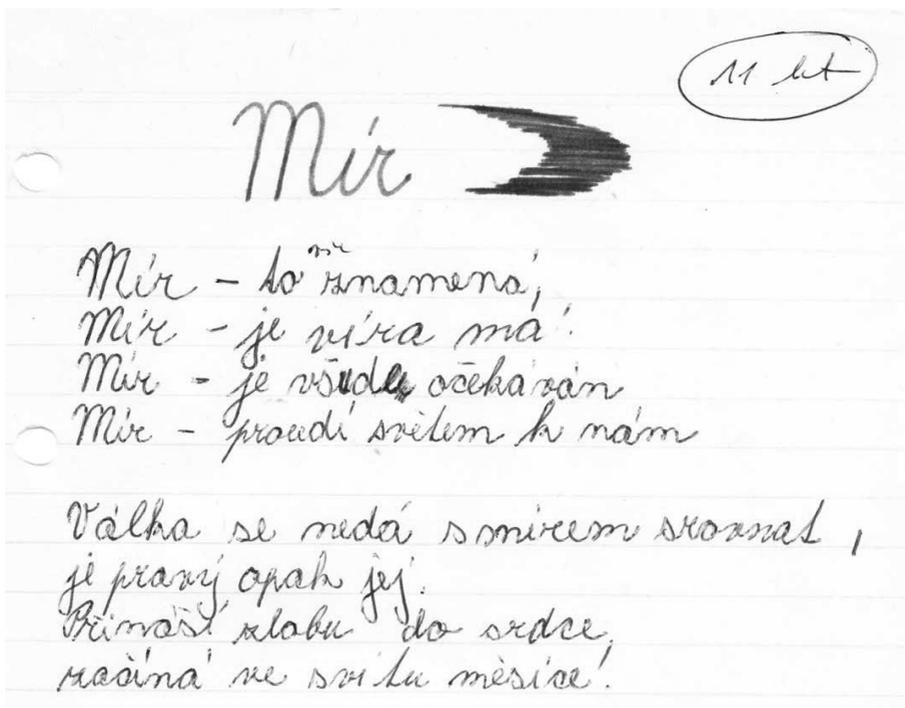


4.1 Evaluation of Transcriptions

Of the seven texts used, we selected three of the most relevant examples for presentation in this study, namely Texts Nos. 1, 2, and 6. The remaining texts are not included here for the following reasons. Sample No. 5, written in block letters, confirmed that there are no significant problems in transcribing such texts. Sample No. 3 – combining block and cursive script – appeared interesting for examination; in this case, the AI performed very well. Similarly, Text No. 4, written in the

same script, was transcribed almost without errors. This type of handwriting did not pose a problem for the AI. Text No. 7 by the second author, like the result presented here for Text No. 6, was transcribed with a high degree of inaccuracy.

Below each text that was examined, the transcription produced by the AI tool is always presented on the left, while the transcription by the authors is shown on the right. The original wording of the texts is in the Czech language and therefore remains preserved here as well. Incorrect transcriptions are highlighted in yellow.



Text No. 1



Chat GPT	Original transcription by the authors
Mír – to znamená, Mír – je víra má. Mír – je <i>vděle</i> očekáván Mír – proudí světem k nám. Válka se nedá s mírem srovnat, je pravý opak <i>jí</i> . Přináší zlobu do srdce, <i>radina</i> ne svitu měsíce!	Mír Mír – to znamená Mír – je víra má Mír – je všude očekáván Mír – proudí světem k nám Válka se nedá s mírem srovnat Je pravý opak jej. Přináší zlobu do srdce, Začíná ve svitu měsíce

Only four words were transcribed incorrectly; they are marked using *italics*.

Claude	Original transcription by the authors
Mír Mír – to znamená, Mír – je víra má! Mír – je <i>vroucí očekávání</i> Mír – proudí světem k nám Válka <i>si</i> nedá s mírem <i>svazat</i> , je pravý opak jej. Přináší <i>slabé</i> do srdce, <i>radina</i> se svitu měsíce!	Mír Mír – to znamená Mír – je víra má Mír – je všude očekáván Mír – proudí světem k nám Válka se nedá s mírem srovnat Je pravý opak jej. Přináší zlobu do srdce, Začíná ve svitu měsíce

Claude handled the text in a similar manner to ChatGPT, with six incorrect transcriptions.

Copilot	Original transcription by the authors
Mír – <i>do smazání</i> ; Mír – <i>na smazání</i> ; Mír – <i>pro smazání</i> ; Mír – <i>když svítá mám jít spát</i> <i>Věčný oheň mého srdce pálí, jako by jsi mi ho do prsou vrazil ostré, mrázivé broty ledu.</i>	Mír Mír – to znamená Mír – je víra má Mír – je všude očekáván Mír – proudí světem k nám Válka se nedá s mírem srovnat Je pravý opak jej. Přináší zlobu do srdce, Začíná ve svitu měsíce

Unfortunately, **Copilot** was essentially unable to transcribe this text; it managed to identify correctly only the word Mír, in which it changed the long *í* to a short one. In effect, it produced an entirely new text.



Šla st'ka' doh' fo troci'
Šluna nede mnou puseh' kocu'
a ja' smat'el a d'ubi' m'aim
a tu' si' pro' s'le' p'rom'it'aim
se p'pam'ln'ka'ch si' p'ri'ad'it' de'la'im
s'y't'm'aj'si' komoru fo d'laubi' doh'i
at'ru'aim
u'chost mi' s'ro'lee' s'p'ro'it'
komoru se at'ru'ka'
a ja' kas' p'p'le'm' t'rou' t'ubi' j'ado' s'el'rae
s'y's't'm' t'rou' s'om'ich i' p'lo'it'
i' j'ad' si' s'y't'al: ~~mo'~~ l'ar'ko' me'
i' d'us' u'it'm' t'rou' d'el'
m'im'aim t'ri' p'ls'ni', t'ri' s'y'p'ad'ni'
a tu' si' p'p'ad'ni'
s'y'p'ad'ni' d'us'e' me', s'y'p'ad'ni'
s'y'd'yt' on' s'ipe' d'ol' a' s'y'ch
p'p'am'ln'ka'ch.

T'm' p'ls'ni' j'ame' tu' d'ol'
a' ja' s' d'ep'it'm'
p'p'ad'ni'

Text No. 2



ChatGPT Slza stéká dolů po tváři Slunce moje mnoze pust' loďí roří a já zmačkl o duši svou a tady si prosím tě přemítám se připomínáš v pořádku dělení pýtají máš komoru po dlouhé době astruim mlhost mi srdce svého! komora se otrává a já zas pěkně tone fráz jak selva svýtm tvým smychem a plóř "Jár si říkal: má" láska mé i dnes cítím tvůj dech mám mám to jsem trá vypráno a teď si žalmi Myšlenky duše mé, myšlenky UNWRITTEN ² přítím on zije dál v mých vzpomínkách. Tvé písně jsou tu dál a já s dojetím poslouchám.	Original transcription by the authors Slza stéká dolů po tváři Slunce nade mnou vesele září A já smutek v duši mám A tak si pro sebe promítám Ve vzpomínkách si pořádek dělám Nejtemnější komoru po dlouhé době otvírám Úzkost mi srdce svírá A já zas vidím tvou tvář jako živou Slyším tvůj smích i pláč I jak si šeptal: láska má I dnes cítím tvůj dech Vnímám tvé písně, tvá vyznání A tak si říkám Neplakej duše má, neplakej Vždyť on žije dál v mých vzpomínkách Tvé písně jsou tu dál A jí s dojetím poslouchám
--	--

This manuscript proved to be somewhat more challenging for **ChatGPT**; the transcription contains a significantly higher number of incorrectly transcribed words, and in some cases entirely new words were generated. In this case, the errors are not marked, as they occur in large numbers.

Claude Sbe stihá' dolů po noci Slunce nedá mnoho pusklé zvoní A já smutně v duši mám A táž si pro sebe znamenám ve vzpomínkách v příšle době nejkrásnější komnatu po dlouhé době otvírám Úzkost mi srdce svírá komnata se otvírá a já zas svým tuším tebou jako zdávna slyším tvůj smích a pláč "již si vytrh' mou lásku mou" i dnes cítím tvůj dech mírním tvé jásní tvá vyznání a těž si písniím Mylenky duše má, mylenky stále on zúje noc v mých vzpomínkách. Tou písni jsem ty, noc a já s během poslouchám	Original transcription by the authors Slza stéká dolů po tváři Slunce nade mnou vesele září A já smutek v duši mám A tak si pro sebe promítám Ve vzpomínkách si pořádek dělám Nejtemnější komoru po dlouhé době otvírám Úzkost mi srdce svírá A já zas vidím tvou tvář jako živou Slyším tvůj smích i pláč I jak si šeptal: láska má I dnes cítím tvůj dech Vnímám tvé písně, tvá vyznání A tak si říkám Neplakej duše má, neplakej Vždyť on žije dál v mých vzpomínkách Tvé písně jsou tu dál A já s dojetím poslouchám
---	---

Claude handled Text No. 2 just as poorly as **ChatGPT**; its transcription also contains a large number of errors.

² At this point in the AI transcription, the word UNWRITTEN appeared.



<p>Copilot She stile doli po trai Chance nade mnou muscle cari a in satis a dute mam a tur si pronte promitam se aspaminhail je poradio dileme atrimam uchoost mi srdce mere homara se atevirá a ja Das sitime tome trace jado silva skyrim Troj somich , flor I por si syptal: banho me i does itini Frag dich mimome thi fissi, trá vyrendbe a ver si resami Mybaby dure mo, mplosei videst on lige del 1 meich KjamiIwhich Tu passé por tu dal a la I depetim</p>	<p>Original transcription by the authors Slza stéká dolú po tváři Slunce nade mnou vesele září A já smutek v duši mám A tak si pro sebe promítám Ve vzpomínkách si pořádek dělám Nejtemnější komoru po dlouhé době otvírám Úzkost mi srdce svírá A já zas vidím tvou tvář jako živou Slyším tvůj smích i pláč I jak si šeptal: láska má I dnes cítím tvůj dech Vnímám tvé písně, tvá vyznání A tak si říkám Neplakej duše má, neplakej Vždyť on žije dál v mých vzpomínkách</p> <p>Tvé písně jsou tu dál A jí s dojetím poslouchám</p>
--	--

Copilot was unable to transcribe the text at all, and it can be stated with some exaggeration that it almost created a new language.

Table 1 Comparative overview of three artificial intelligence tools (ChatGPT, Claude, and Copilot) based on their performance in the transcription and analysis of historical handwritten texts.

AI Tool	Accuracy of transcription of the poem written by a child	Accuracy of transcription of more complex handwritten texts	Quality of text analysis	Ability to detect the writing instrument	Sensitivity to free form prompting	Overall evaluation
Chat-GPT	High	Low	Medium	Blue ink (incorrect)	Good	2nd place
Claude	High	Low	High	Blue ink (incorrect)	Good	1st place
Copilot	Low (completely incorrect transcription)	None	Low/none	Blue ink (incorrect)	Low (misunderstanding)	3rd place



AKTA 4

S: maldři 2x psály: m. m.

M: Defej S. Tady v kade uží No
mimí lyž.

S: J. No rovnýd M. Proč myslíš
že s duchem není o nějakých
mimí kumírka. Proč už nevíš?

M: Ne S. ^{Angličtina} Informace o nich jsem
najel. Tady. Tadyji Anka, tu
Armada převedla podruky k yobi
Měsíček obléh? pro malomocnost.
EM.

S: m. Popale nad jízdem,
udil ročník pro jízdem

M: pro EM dle

S: Eje, Eje m. lochy
do ^{Angličtina} ~~Angličtina~~ a do rouchy.
All to nic nemamem M.

M: ne? a od Nohe. Dnes je



<p>ChatGPT AKTA 4 S: Maléři se zapíší na mě. M: Díky S. Tady přidávám ještě to menu týká. S: Je to zapsáno M. Ono myslíš. Ptá po avých dnech není o podmínkách mých pomáhá. Pokud nemyslíš, zapíší vynechávat kryty. M: Ne S. Informace o mých jsou raje. Tady, Trojský Arihack, při armádách pamatují návrhy k rozšíření švihátek obleků pro uměním úct. S: Na řádku pod jménem, rodič zakni pod večerem. M: Pro čas díl. S: Díky, díky za úkoly, a měřit. Alt-moc chyby, ale to nic neznamená M. M: Ne? A od Noble. Dnes je.</p>	<p>Original transcription by the authors Akta Y S: Moldře 2x počkej na mě. M: Dělejš S. Tady někde už to musí být. S: Je to nesmysl M. Proč myslíš, že v archivech není o vodnicích ani zmínka. Protože neexistují. M: ne S. Protože informace o nich jsou tajné. Tady. Tady je důkaz, že armáda zneužila vodníky k výrově speciálních obleků pro námořnictvo. Čti. S: Na topole nad jezerem seděl vodník pod večerem. M: No čti dál. S: Šiju, šiju si botičky do sucha i do vodičky. Ale to nic neznamená M. M: Ne? A co tohle. Dnes je</p>
--	---

ChatGPT was essentially unable to transcribe it; this handwriting was very complex.

<p>Claude AKTA 4 S: Prosím, dr. podpis: M. mi. M: Dějí S. Tady nikdo má to mim byt. S: Já to chápal M. Brož myslí, že s archivech není a nivíte mi pamětin. Brožův neúspěch M: Ne S. CHYBÍ Informace o mně jsou tajné. Tady. Tady je Archiv, tři armáda speciální oddělení h syni- speciálních oblastí pro mimořádnosti. Čti. S: Jsi několi pod jedem, abys mluvil pod názvem M: No evi dál S: Říjen, říjen na dobytky dr. podpis n dr. podpisy. Ale to jsou neznámé M. M: Ne? A co hodle. Dnes je</p>	<p>Original transcription by the authors Akta Y S: Moldře 2x počkej na mě. M: Dělejš S. Tady někde už to musí být. S: Je to nesmysl M. Proč myslíš, že v archivech není o vodnicích ani zmínka. Protože neexistují. M: ne S. Protože informace o nich jsou tajné. Tady. Tady je důkaz, že armáda zneužila vodníky k výrově speciálních obleků pro námořnictvo. Čti. S: Na topole nad jezerem seděl vodník pod večerem. M. No čti dál. S: Šiju, šiju si botičky do sucha i do vodičky. Ale to nic neznamená M. M: Ne? A co tohle. Dnes je</p>
--	---

Claude transcribed the text with numerous errors.



<p>Copilot Bohužel nemohu provést doslovný přepis textu z obrázku, protože můj současný systém mi neumožňuje číst dlouhé texty. Pokud potřebujete pomoc s krátkými texty nebo máte další dotazy, dejte mi prosím vědět!</p>	<p>Original transcription by the authors Akta Y S: Moldře 2x počkej na mě. M: Dělej S. Tady někde už to musí být. S: Je to nesmysl M. Proč myslíš, že v archivech není o vodnících ani zmínka. Protože neexistují. M: ne S. Protože informace o nich jsou tajné. Tady. Tady je důkaz, že armáda zneužila vodníky k výrobě speciálních obleků pro námořnictvo. Čti. S: Na topole nad jezerem seděl vodník pod večerem. M. No čti dál. S: Šiju, šiju si botičky do sucha i do vodičky. Ale to nic neznamená M. M: Ne? A co tohle. Dnes je</p>
--	--

Copilot did not transcribe this text, stating that it was unable to read long texts.

Across all the tools, the same error occurred – the writing instrument was described as blue ink, although in reality it was a blue ballpoint pen. This indicates certain limitations in the visual differentiation of commonly-used writing instruments at the end of the twentieth century. In most cases, none of the transcriptions produced by any of the tools would be fully usable at 100% accuracy, certainly not without further revision and ideally not without prior knowledge of the transcribed source for verification purposes.

The indicative quantification of transcription accuracy revealed significant differences between the individual tools. For the poem *Mír* (Text No. 1), the ChatGPT model achieved approximately 90% accuracy in transcribing words correctly, while Claude achieved around 85%. At the same time, Copilot reached only single-digit percentages (with most

of the text replaced by newly-generated content). For the more demanding handwritten text with a love theme (Texts Nos. 2 and 6), accuracy dropped substantially for both models – Claude achieved an estimated accuracy of around 25% correctly transcribed words and ChatGPT approximately 20%, while Copilot again generated largely unintelligible or entirely new text, resulting in a realistically usable accuracy well below 5%. These indicative figures confirm that even in cases where a transcription appears “*partially readable*” at first glance, its reliability for research purposes remains limited and requires thorough human verification.

4.2 Analysis of Transcribed Texts Using AI

In addition to transcribing handwritten texts, another task was to determine



the ability of the artificial intelligence tools to analyse and interpret the text. This step aimed to verify whether the individual models were capable of recognising not only the content and formal features, but also the semantic and stylistic layers of the documents that were processed. The intention was to simulate a situation in which AI serves not only as a tool for the technical processing of handwriting, but also as a potential collaborator in scholarly research.

The analysis focused on several key aspects:

- Thematic classification of the text, that is, the ability of AI to determine what the text is about.
- Emotional tone, i.e. whether AI can identify the feelings and moods expressed or evoked by the text.
- Stylistic characteristics, particularly the distinction between different types of texts (poetry vs. notes, childlike vs. adult style).
- Interpretative ability, that is, whether the model offers an interpretation of the content or at least basic contextualisation.
- Cultural and historical anchoring, where such elements were present in the outputs.

In terms of output quality within this limited sample, the Claude model performed best. It not only handled transcription with the highest accuracy but also attempted a brief interpretation in several cases, suggesting a possible theme or emotional dimension of the text. ChatGPT produced less consistent interpretations; however, in some cases, it captured the overall mood of the poem and

responded in line with the given prompt. By contrast, Copilot failed in most cases – not only in transcription, but also in subsequent work with the content – as it was unable to grasp the meaning of the texts and often generated irrelevant or superficial outputs.

The results indicate that the ability to analyse and interpret texts represents a more demanding task for the generative models that were tested than transcription itself. Nevertheless, their potential cannot be overlooked, particularly in the role of an assistant that can offer an initial interpretative outline, formulate alternative perspectives on the text, or guide the user towards deeper reflection on its content. Such use, however, always presupposes critical oversight and final evaluation by a human – whether a researcher, archivist, or educator.

4.3 Case Study – Comparison of the Analysis of the Poem “*Slza stéká dolů po tváři*”

As a model example for comparing the ability of AI to analyse handwritten texts, a poem from the transition between adolescence and adulthood was selected. Both in terms of content and language, it carries a strong emotional charge and complex metaphorical layers.

In its analysis, **ChatGPT** recognised that the text is a lyrical poem with a distinctly melancholic tone. It identified words such as “tear”, “sadness”, and “memories” and interpreted them correctly as indicators of loss, nostalgia, and



emotional pain. It further characterised the structure as free verse and described the text as an “introspective and subjective statement”. Although some expressions were already incorrectly decoded at the transcription stage, the interpretation appeared coherent and sensitive.

Sample of part of the interpretation by ChatGPT³

Style: analytical, empathetic, linguistically rich

“The poem ‘Slza stéká dolů po tváři’ can be read as an introspective lyrical statement that thematises inner pain and reconciliation. The author employs metaphorical language and builds contrasts between light and darkness, and between consciousness and memory. What is particularly notable is the repeated motif of song as a carrier of memory, which highlights the link between the past and the speaker’s present emotional experience. From a palaeographic perspective, the writing alternates between cursive and block capitals, which may reflect emotional tension during the act of writing...”

Claude offered the most comprehensive interpretation. It not only recognised the poetic nature of the text, but also attempted to situate it within a broader context (e.g. a “reflection on the loss of a loved one”). It identified figurative devices (verbal metaphors, synecdoches) and pointed out linguistic features such as parallelism and anaphora. In the conclusion of the

analysis, it stated that the poem “exhibits signs of emotional struggle, personal grief, and inner catharsis”.

Sample of part of the interpretation by Claude

Style: factual, logical, at times almost overly “academic and formulaic”

“The poem describes the emotional experience of a subject who reflects on a past relationship. The text contains metaphors related to pain, loss, and memories. The author employs contrasts between light and darkness to express psychological tension. The poem is written in free verse. On the basis of the analysis, it can be assumed that the handwritten material dates from the late twentieth century, which corresponds with the use of common school paper and ink.”

By contrast, **Copilot** did not provide any meaningful analysis. As a result of the incorrect transcription of the text, it generated content that bore no relation to the original text. The attempt at interpretation was based on artificially generated sentences that lacked profound meaning and were not grounded in the actual linguistic material.

Sample of part of the interpretation by Copilot

Style: hesitant, lacking context, low interpretative ability. If it does not understand the text, it often “hedges”.

³ All prompts and model outputs were originally produced in Czech. For the purposes of this manuscript, excerpts are presented in English as author translations. Translations aim to preserve meaning rather than literal phrasing; minor stylistic smoothing in English does not affect the evaluation, which was conducted on the Czech originals.



“This text appears to be a mixture of different languages and words, which may be difficult to interpret. Some parts seem to be English, others Czech, and still others a completely different language or even nonsensical combinations. For example, ‘She stile doli po trai’ may be an attempt at an English sentence, but it does not make sense; ‘Chance nade mnou muscle cari’ contains Czech words but also lacks clear meaning; and ‘Mybaby dure mo, mplosei’ looks like English words mixed with something else. Overall, it appears to be a text that was either poorly translated, or an experimental poem or artistic text intended to evoke certain feelings or images rather than a specific meaning.”

This case study demonstrates that current AI models, when properly guided, can provide analytical outputs that complement classical humanities analysis. The greatest weakness remains their dependence on the quality of transcription; incorrect word identification has a fundamental effect on the overall interpretation. Nevertheless, the potential of AI can be recognised in the role of an assistant for the basic structuring and annotation of handwritten texts.

5. ETHICAL ASPECTS OF THE RESEARCH

The use of artificial intelligence in the transcription and analysis of, among others, historical and handwritten texts brings several advantages, but it also raises fundamental ethical issues that must be addressed. When working with handwrit-

ten materials – whether real or synthetically generated – it is essential to emphasise the transparency of the methods used, a conscious awareness of the limitations of the tools, and responsibility for interpreting the results (Philips & Tabrizi, 2020). Language models, for example, may “fill in” missing parts or infer semantic connections when interpreting a historical manuscript, which can lead to an inevitable contamination of the original message (Celli & Spathulas, 2025).

One of the key ethical aspects concerns the issue of the authenticity and trustworthiness of the outputs generated by AI. Although tools such as ChatGPT or Claude are capable of offering interesting and, in some instances, convincing analyses, it is essential to bear in mind that they are not human researchers and therefore cannot be held responsible for potential distortions or errors in interpretation. Responsibility for the manner in which these tools are used, as well as for the resulting interpretation, always lies with the researcher, archivist, or educator. This also applies to the work of students, who may use these tools without a thorough knowledge of historical context and methodology, which can lead to incorrect or oversimplified conclusions (Epstein et al., 2023).

It is also worth mentioning the issue of equal access to technology. Not every student or researcher has access to paid versions of AI tools or to high-performance hardware, which may deepen existing inequalities in education and research. In our approach, we therefore used



free versions of the tools and deliberately simulated prompts formulated in a looser, “informal” style that would be likely to be used by an ordinary user without more profound knowledge of prompt engineering – for example, a student who needs to transcribe palaeography notes, or an instructor testing the accessibility of technologies for their students (Crosilla et al., 2025).

Finally, it is also necessary to address the issues of personal data protection and copyright if real handwritten materials are to be processed (e.g. student assignments, family letters, or unpublished archival documents). In such cases, it is essential to consider the sensitivity of personal data, obtain informed consent, anonymise texts where appropriate, and respect copyright and licensing restrictions in all cases (Lauricella & Kay, 2013). In our research, only authorial texts from the personal archive of one of the authors were used for experimental purposes, and these were not shared beyond the scope of the research.

Attention should also be paid to the issue of epistemic security – that is, the degree to which AI outputs are regarded as reliable and trustworthy. Given that language models generate text on the basis of probabilistic patterns rather than human understanding, their outputs may not always be truthful, consistent, or methodologically grounded (Bender et al., 2021). It is therefore essential that users of AI tools – especially students or early-career researchers – are guided towards careful verification, critical read-

ing, and reflection on the results, rather than towards their uncritical acceptance (Carey, 2023).

From an ethical perspective, AI can therefore be viewed primarily as a supportive tool that streamlines technical steps (such as transcription) and provides an initial interpretative outline, rather than as an authoritative source of historical knowledge. The key role must continue to be played by the human expert, with their professional judgement and methodological rigour.

6. EDUCATIONAL BENEFITS OF AI TECHNOLOGIES IN THE FIELDS OF HISTORICAL STUDIES AND ARCHIVAL SCIENCE

Artificial intelligence presents new opportunities to expand access to historical documents, not only for scholars but also for students and the broader public. Automated transcription systems and text recognition tools enable access to handwritten documents for individuals without palaeographic training, thereby significantly increasing the accessibility of cultural heritage (Philips & Tabrizi, 2020). In the context of this article, a similar approach can be applied, for example, when working with students’ palaeography notes or practice manuscripts, where AI provides an initial transcription. In contrast, the student compares the output with the original and learns to identify errors and the limitations of the technology. Tools such



as ChatGPT, Claude, or Copilot may, in some instances,⁴ serve as assistants for self-study, helping users explore the context of historical sources, search for related information, or verify their own transcriptions.

In this way, the concept of “learning with AI”, in which artificial intelligence is not only the subject of instruction (i.e. an object of critical reflection), but also a tool that supports the educational process, is strengthened (Godwin-Jones, 2024). In the teaching of historical studies or archival science, this may involve, for example, assignments in which students work with AI-generated outputs as a “hypothesis” that they are required to verify on the basis of primary sources and scholarly literature.

In the field of archival science, AI can further assist in the digitisation, classification, and description of archival fonds, and in some cases, also in predicting the content of as-yet-undigitised documents on the basis of metadata. Such applications can contribute not only to the efficiency of archivists’ work but also to a more systematic approach to historical information (Emerson et al., 2020). At the same time, new didactic methods are emerging that involve students in practical tasks, such as designing descriptive fields, verifying

automatically generated descriptions, or critically evaluating the relevance of AI outputs (Mah & Gross, 2024).

At the same time, new didactic approaches are emerging that integrate AI directly into teaching as a tool to support critical thinking, work with errors, and interdisciplinary reasoning (Mah & Gross, 2024). For example, instructors may ask students to compare transcriptions and interpretations generated by different AI models, identify inaccuracies, and discuss the potential consequences of uncritically adopting such outputs for historical research or archival practice. In combination with the ethical reflection described in the previous section AI can thus serve as a powerful didactic tool – however, only when it is used consciously, transparently, and with an emphasis on the responsibility of the human user.

CONCLUSION

This research study demonstrated that generative artificial intelligence has considerable potential in the transcription and interpretation of handwritten texts, while also revealing its current limitations. On the basis of the limited corpus of modern personal handwritten texts tested in this study, the Claude model developed by

⁴ We recommend that all information generated by artificial intelligence be verified with regard to its factual accuracy, relevance, and contextual appropriateness. Especially in academic or educational settings, it is advisable to consider the use of paid versions of generative AI systems (e.g. ChatGPT Plus or higher-tier versions), which typically provide access to a broader range of models and demonstrate a higher level of reliability and accuracy. Furthermore, it is beneficial to make use of the option to upload one’s own expert materials into the environment of a selected AI tool, as this can significantly enhance the relevance and contextual grounding of the outputs that are generated.



Anthropic proved to be the most effective of the three tools, excelling not only in text transcription but also in fundamental analysis and interpretation. ChatGPT by OpenAI demonstrated a good level of understanding and, in many cases, provided reasonably high-quality analyses, whereas Microsoft's Copilot lagged significantly behind in the tasks that were examined. A common shortcoming of all the models was, for example, the incorrect identification of the writing instrument (blue ink instead of a ballpoint pen), which highlights the need for further development of visual detection capabilities and more precise differentiation of details when working with documents.

From the perspective of transcription, the versions of the tools that were tested (primarily from late 2024) appeared to be usable mainly for block-letter texts or simpler handwriting, while more complex and individually formed scripts, as well as a higher degree of variability, posed significant difficulties for the systems. Transcription accuracy directly affects the quality of subsequent interpretation, which confirms that both phases – the technical phase (transcription) and the analytical phase (interpretation) – are closely interconnected in the case of textual sources. The indicative quantification of errors further showed that even for relatively simple handwriting, the performance of the tools differs markedly (with Claude and ChatGPT achieving high but not perfect accuracy, while Copilot fails almost entirely), and that for more complex texts, the proportion of correct-

ly transcribed words drops to a level that fundamentally limits the possibility of reliable interpretation without expert intervention.

Another significant benefit is the possibility of using these tools in education. From a didactic perspective, the most critical potential of these tools can be seen particularly in the teaching of palaeography, modern diplomatics, archival science, and digital humanities, where they can facilitate students' engagement with handwritten texts without requiring them to master all the specialist skills immediately. At the same time, they enable the verification of AI outputs, thereby strengthening critical thinking. Artificial intelligence enables students and the wider public to access texts that would otherwise be difficult to read or inaccessible without palaeographic or diplomatic training. This can support individualised learning, foster interest in cultural heritage, and contribute to the development of new teaching methods that bridge the humanities and technological disciplines. When combined with critical reflection on outputs and an emphasis on ethical aspects, AI can become a valuable didactic tool, for example – as noted earlier – in the fields of palaeography, archival science, and digital humanities.

The results simultaneously confirm the need for a cautious and critical approach to interpreting outputs generated by generative AI from both methodological and ethical perspectives. Preserving the integrity of documents, ensuring transparency in the use of tools, taking



due account of the protection of personal data, and consistently verifying outputs should be integral components of any work that employs these technologies. In historical scholarship, AI should therefore not be understood as a replacement for human expertise, but rather as a potentially valuable complement – a tool for accelerating, extending, and in some instances refining the research and educational process.

This research also opens the way for further investigation, particularly in the application of generative artificial intelligence to older historical texts, for example, sources written in Kurrent, humanistic cursive, or other older types of script. Although specialised systems such as Transkribus AI,⁵ PERO OCR,⁶ and other tools employing HTR/OCR technologies already exist, we believe that generative lan-

guage models such as ChatGPT, Claude, and Copilot (and others) may play an essential role in the future – especially in the areas of educational support, self-directed learning, and preliminary text analysis. Their ability to combine technical transcription with linguistic analysis and content interpretation could offer a more accessible and user-friendly alternative for both students and the broader scholarly and general public. With the continued development of these models, opportunities are emerging for interdisciplinary experiments that may enrich not only humanities research but also the didactics of historical studies and archival science. It is precisely in the intersection of history, archival studies, computer science, and the development of artificial intelligence that we see one of the key challenges and opportunities for the future.

REFERENCES

- ABBYY. (2023). *Old fonts recognition*. <https://pdf.abbyy.com/learning-center/old-fonts-recognition>
- Alkendi, W., Gechter, F., Heyberger, L., & Guyeux, C. (2024). Advancements and challenges in handwritten text recognition: A comprehensive survey. *Journal of Imaging*, 10(1), Article 18.
- Arthur, K., Müller, R., & Strobel, H. (2004). Recognizing digitization as a preservation reformatting method. *Microform & Imaging Review*, 33(4), 171–177.
- Baheti, P. (2022). Handwriting recognition: definition, techniques & uses. *V7 Labs*. <https://www.v7labs.com/blog/handwriting-recognition-guide>
- Bartz, C., Seidel, L., Nguyen, D.-H., Bethge, J., Yang, H., & Meinel, C. (2020). Synthetic data for the analysis of archival documents: Handwriting determination. *2020 Digital Image Computing: Techniques and Applications (DICTA)*, Melbourne, Australia, 1–8.

⁵ <https://transkribus.ai>

⁶ <https://pero-ocr.fit.vutbr.cz>



- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.
- Berchmans, D., & Kumar, S. S. (2014, July). Optical character recognition: An overview and an insight. *2014 International Conference on Communication and Network Technologies (ICCNT)*, 59–63.
- Carey, E. (2023). *AI (artificial intelligence) in education: Evaluating sources*. <https://libguides.sbcc.edu/c.php?g=1306678&p=10027413>
- Celli, F., & Spathulas, G. (2025). Language models reach higher agreement than humans in historical interpretation. *arXiv:2504.02572*.
- Colavizza, G. et al. (2021) Archives and AI: An overview of current debates and future perspectives. *Association for Computing Machinery*, 15(1), 1–15.
- Crosilla, G., Klic, L., & Colavizza, G. (2025). Benchmarking large language models for handwritten text recognition. *arXiv:2503.15195*.
- de Sousa Neto, A. F., Bezerra, B. L. D., de Moura, G. C. D. et al. (2024). Data augmentation for offline handwritten text recognition: A systematic literature review. *SN Computer Science*, 5, 258.
- Dietrich, F. (2021, May 9). *OCR vs. HTR or “What is AI, actually?”*. READ-COOP. <https://readcoop.eu/insights/ocr-vs-htr/>
- Emerson, A., Cloude, E. B., Azevedo, R., & Lester, J. (2020). Multimodal learning analytics for game-based learning. *British Journal of Educational Technology*, 51(5), 1505–1526.
- Epstein, Z., Hertzmann, A., Herman, L. et al. (2023). Art and the science of generative AI: A deeper dive. *arXiv:2306.03819*.
- Ferro, S., Pelillo, M., & Traviglia, A. (2023, June 24). AI-assisted digitalisation of historical documents. *ISPRS, XLVIII-M-2-2023*, 557–563.
- Frontoni, E., Contigiani, M., Zingaretti, P., Bernardini, M., Placidi, V., & Mecocci, A. (2022). Trusted data forever: Is AI the answer? *arXiv:2203.03712*.
- Gartner. (2024). *Gartner experts answer the top generative AI questions for your enterprise*. <https://www.gartner.com/en/topics/generative-ai>
- Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. *arXiv:2403.20216*.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M. et al. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139–144.
- Gozalo-Brizuela, R., & Garrido-Merchán, C. E. (2023). A survey of generative AI applications. *arXiv:2306.02781*.
- Hedau, S. (2024). *OCR vs ICR: Battle of digitizing printed & handwritten text*. *Softspace Solutions*. <https://softspacesolutions.com/blog/ocr-vs-icr/>
- Kadaruddin, K. (2023). Empowering education through generative AI: Innovative instructional strategies for tomorrow’s learners. *International Journal of Business, Law and Education*, 4(2), 618–625.



- Lauricella, S., & Kay, R. (2013). Exploring the use of text and instant messaging in higher education classrooms. *Research in Learning Technology*, 21, 19061.
- Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education*, 22(1), 14.
- Leviner, S. (2023). The role of OCR in digitizing historical and archival documents. *CharacTell*. <https://www.charactell.com/resources/the-role-of-ocr-in-digitizing-historical-and-archival-documents/>
- Ma, Y., Liu, J., & Fan, Y. (2023). AI vs. human: Differentiation analysis of scientific content generation. *arXiv:2301.10416*.
- Mah, D.-K., & Gross, N. (2024). Artificial intelligence in higher education: Exploring faculty use, self-efficacy, distinct profiles, and professional development needs. *International Journal of Educational Technology in Higher Education*, 21(1), 58.
- Mello, F. R., Lopes, R. R., Oliveira, R. D., & Silva, A. P. (2023). Education in the age of generative AI: Context and recent developments. *arXiv:2309.12332*.
- Memon, J., Sami, M., Khan, R. A., & Uddin, M. (2020). Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). *IEEE Access*, 8, 142929–142951.
- Mondal, A., & Jawahar, C. V. (2022). Enhancing Indic handwritten text recognition using global semantic information. *arXiv:2212.07776*.
- Nockels, J., Gooding, P., Ames, S. et al. (2022). Understanding the application of handwritten text recognition technology in heritage contexts: A systematic review of Transkribus in published research. *Archival Science*, 22, 367–392.
- Philips, J., & Tabrizi, N. (2020). Historical document processing: A survey of techniques, tools, and trends. *arXiv:2002.06300*.
- Sabharwal, A. (2015). Digital history, archives, and curating digital cultural heritage. In A. Sabharwal (Ed.), *Digital curation in the digital humanities* (pp. 33–56). Chandos.
- Spina, S. (2023). Artificial intelligence in archival and historical scholarship workflow: HTS and ChatGPT. *arXiv:2308.02044*.

Klára Rybenská

*University of Hradec Králové, Philosophical Faculty;
email: klara.rybenska@uhk.cz*

Mgr. Sylva Sklenářová, Ph.D.

*University of Hradec Králové, Philosophical Faculty;
email: sylva.sklenarova@uhk.cz*

RYBENSKÁ, K., SKLENÁŘOVÁ, S. Využití generativní umělé inteligence při digitalizaci tištěných a rukopisných dokumentů a její přínos pro historické a archivní vzdělávání

Cíle: Studie zkoumá, jak mohou současné generativní jazykové modely podporovat archivní a historickou práci s českými rukopisnými texty, se zaměřením na transkripci a základní předběžnou analýzu, a vymezuje klíčová omezení i etické souvislosti jejich využití ve výuce v archivnictví a digitálních humanitních vědách.

Metody: Byla provedena kvalitativní případová studie na souboru sedmi moderních osobních rukopisných českých textů z 80. a 90. let 20. století (báseň psaná dítětem, milostné básně, školní test a studijní poznámky). Tři široce dostupné nástroje v bezplatných verzích (ChatGPT, Claude a Copilot) byly testovány za použití shodných instrukcí k úloze. Výstupy byly komparativně hodnoceny z hlediska přesnosti transkripce, obsahové a stylistické interpretace a rozpoznání vybraných formálních rysů textů. Empirické srovnání bylo doplněno kritickou analýzou relevantní odborné literatury a reflexí otázek autenticity, integrity dat, epistemické bezpečnosti a ochrany osobních údajů.

Výsledky: Nejlepší celkový výkon dosáhl model Claude, následovaný ChatGPT, zatímco Copilot vykázal v testovaných úlohách výrazně slabší výsledky. Napříč nástroji se ukázalo, že interpretace a analýza jsou pro generativní modely náročnější než samotná transkripce; výstupy navíc obsahovaly chyby a nadinterpretace, které vyžadovaly odborné ověření.

Závěry: Generativní jazykové modely mohou fungovat jako podpůrné nástroje pro transkripci, předběžnou analýzu a didaktickou práci, nemohou však dosud nahradit profesionální archivní ani historickou expertizu. Odpovědné využití vyžaduje kritický lidský dohled a explicitní zohlednění etických a datově-ochranných aspektů.

Klíčová slova: generativní umělá inteligence, velké jazykové modely (LLM), rukopisné texty, HTR (handwritten text recognition), archivnictví, historický výzkum, transkripce rukopisu, etika AI, digital humanities, využití AI ve vzdělávání