OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

The Role of AI in the Decoding of Lost/Ancient Texts

Flueraşu Sara-Alexandra
alexfluerasu06@gmail.com

ABSTRACT

The automated decipherment of lost languages has emerged as a critical frontier in digital humanities, yet the epistemological and technical limitations of AI models remain underexplored and misrepresented. This paper proposes to study 'To what extent can AI language models, such as NLPs trained in low-resource settings, help with the deciphering of lost languages, and what are the actual limitations relative to traditional methods used by linguists?'. It argues that, despite general fear of AI replacement, while AI models accelerate decipherment workflows, they cannot operate independently, as they require expert validation and interpretation to generate reliable results, via a thorough analysis of traditional methods compared to AI-assisted models, backed up by successful case studies such as Linear B and Ugaritic. The results demonstrate how these models offer valuable advantages when working in low-data environments. The research deduces that AI is a tool with immense potential, which can be reached only through a human-machine collaboration that regards the ethical concerns of these processes.

INTRODUCTION

The preservation and understanding of ancient texts is a particularly important factor in acknowledging cultural patrimony. Owing to preservation issues that have resulted in material loss and textual degradation, the process has been impeded. Technological developments in different philological areas inspired researchers to try AI and NLP approaches, which can help overcome the aforementioned obstacles.

This paper proposes to explore the following question: 'To what extent can AI language models, such as NLPs trained in low-resource settings, help with the deciphering of lost languages, and what are the actual limitations relative to traditional methods used by linguists?' The hypothesis it tests is that AI cannot fully replace human expertise, but it can aid experts, resulting in a considerably accelerated process. The paper aims to evaluate the epistemological performances relative to the challenges faced by AI models compared to traditional methods.

There is a long tradition of manual deciphering, and the recent collaboration between these models (such as PYTHIA and NeuroDecipher) has resulted in a shift in efficiency and data interpretation. This study promises to fill the gap in the analysis of efficiency through the lens of interpretability concerns and ethical practices of AI in philology.

BACKGROUND

Ancient languages conserve historical and ethnic sources for the identity of current and extinct civilizations. Consequently, their interpretation is essential since they serve as bonds between cultural heritage and diverse areas of study, such as anthropology and art. Their study is teeming with challenges; however, scholars must overcome impediments such as the restoration of damaged texts and deciphering ancient languages (Sommerschield et al., 2023).

Humans have long understood their importance, and they have laboriously worked on the translation for centuries. Linguists and archeologists manually inspect, methodize, and explain cognates (words that share a common origin), a process that requires attention to detail and time (Diao et al., 2025). Certain extinct languages, such as Linear B, Ugaritic, and Egyptian hieroglyphs, required special attention as their decipherment could lead to understanding key moments in the development of humanity. Thus, the need for digital instruments has emerged. Consequently, automated image processing has become essential to analyzing ancient texts at scale (Diao et al., 2025).

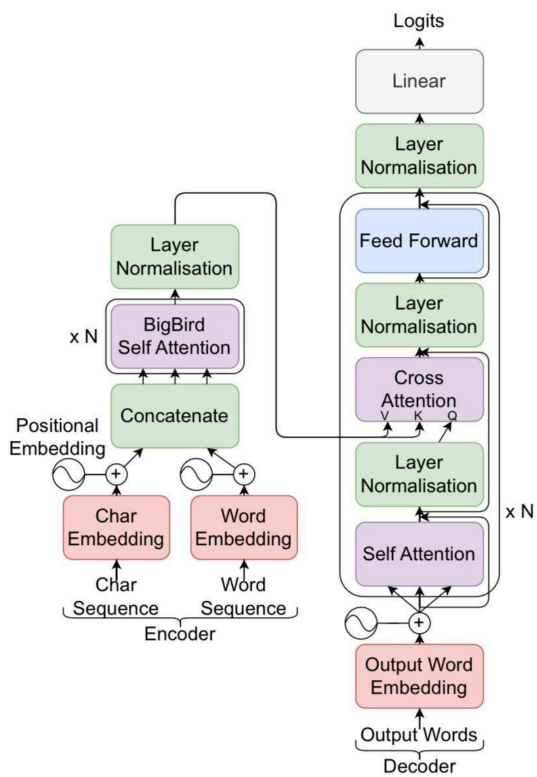
Ancient texts present two interlocking challenges: material degradation (physical degradations occurred over long periods of time, resulting in the appearance of notable real-world noise) (Diao et al., 2025) and linguistic restrictions (scarcity of standardized data due to preservation issues and historical circumstances) (Sommerschield et al., 2023).

Recent works propose a classification system that groups ancient scripts into phonographic (alphabetic or syllabic) and logographic families based on the way linguistic units are encoded. By aligning scripts that share structural properties, researchers can transfer preprocessing pipelines and model architectures across historically unrelated corpora, hence reducing the need for script-specific engineering (Chen, D., Shi, F., Agarwal, A., et al., 2024).

The development of deep learning models has allowed a thorough examination of lexical and phonetic patterns in large corpora of data. As a result, the pattern recognition ability relies heavily on the quality, the quantity of the data, and the evaluation procedure, relative to the metrics used (Sommerschield et al., 2023). This calls for an interdisciplinary approach between AI and anthropology (philology and archeology).

In recent research, the development of AI models dedicated to decipherment has marked a breakthrough in this study area. Their ability to function in low-resource settings is remarkable. This has shifted the focus from automated identification of statistical patterns to applying these processes to the actual decipherment process. The new models were built on a universal principle, offering a more flexible approach, because ancient texts make it close to impossible to apply traditional methods of automated translation, which were based on supervised learning (Luo, Cao & Barzilay, 2019). Additionally, modern models are designed to have a blended architecture, being able to find meaningful connections between words, sentences, and phrases (Luo, Cao & Barzilay, 2019). Key implementations of this approach include Luo, Cao & Barzilay's neural decipherment model (67% accuracy on Linear B), DeepMind's PYTHIA system for Greek inscription restoration, and NeuroDecipher for unsupervised cognate alignment.

figure 1



To improve AI's approach to multiple handwriting styles, newer models use an approach that allows the network to learn and decide which visual features of the handwriting are the most important for text recognition. Combined with an attention mechanism that can identify text lines automatically, the system can transcribe an entire paragraph at once. This avoids a method that isolates the steps, leading to more accurate results (Hamdan & Cheriet, 2023).

To understand how these modern models actually read the texts, the best approach would be to look at the architecture behind them. Figure 1 shows the inner structure of a Transformer model and its core technology. The model takes two kinds of inputs, characters and words, and joins them through embedding layers. After, it uses attention mechanisms (visible in the central blocks of the diagram) to assess which parts of the script are prioritary when looking for a translation. Instead of reading left to right, these models can process the entire sequence at once. Thus, they can easily identify patterns between languages even with scarce and unclean data (Larth: Dataset and Machine Translation for Etruscan).

For example, a targeted HTR model trained on only 3,000 manually transcribed lines reduced the character-error rate from 12% to 5.8% for the Belfort birth registers. Thus, a two stage layout-analysis combined with a relatively scarce amount of ground-truth data can improve accuracy on archives of diverse formats (Plateau–Holleville et al., 2024).

The field has moved from early HMM deep-learning pipelines to operate on whole pages. Modern convolutional-transformer hybrids can jointly learn visual features and language models, enabling character-level transcription without an explicit segmentation stage. This transition has opened the door to large-scale analysis of corpora that were previously unsolved (Coquenet et al., 2021). To avoid the slow sequential processing of recurrent layers used in older models, researchers developed GFCNs:

recurrence-free gated fully convolutional networks. They replace LSTMs with convolutional gates that capture long-range dependencies without hindering parallel computation. As a result, training time can be reduced by up to 45% while keeping the accuracy levels similar to the recurrent models'(Coquenet, Chatelain, & Paquet, 2020).

Contemporary research priorities have shifted toward three areas: multilingual models trained simultaneously on multiple language families, few-shot learning architectures that require minimal training data, and how to facilitate a better environment that embeds expert oversight into the decipherment pipeline.

RISE OF AI AND NLP IN THE DECIPHERMENT PROCESS

1. AI models used in the decipherment of the texts:

Classical decipherment methods impose laborious work, as they consist of linguists manually examining old texts and manuscripts. Instead, AI models approach decipherment as a puzzle, an analytical issue that can be solved via computational algorithms. Recently, there has been a rapid development regarding these procedures, such as improving the algorithms that can identify patterns, compare unknown symbols, and suggest prospective translations and interpretations in the absence of direct human interference.

Commonly used models often vary between unsupervised statistical models, such as conditioned entropy or minimum-cost flow (an optimization technique that enforces one-to-one character mappings between language pairs), and unsupervised neural networks (seq2seq, a neural architecture that takes a sequence of inputs and generates a sequence of outputs), which transform texts into numeric codes. Luo, Cao & Barzilay (2019) developed a model that aligns unknown words from Linear B with their respective equivalent in ancient Greek, having a success rate of roughly 67%. Moreover, Tamburini (2025) proposes an algorithm that uses simulative annealing for global optimization to test multiple hypotheses regarding translations and meanings.

Automated transcription involves several steps to turn manuscript images into machine-readable text. Figure 2 shows the standard HTR pipeline, split into two main parts: training and testing. The training phase uses cleaned images to build a model file. In the testing phase, new manuscript images go through the same steps before the models identifies the text. Each stage is equally important, affecting the result directly (AlKendi et al., 2024).

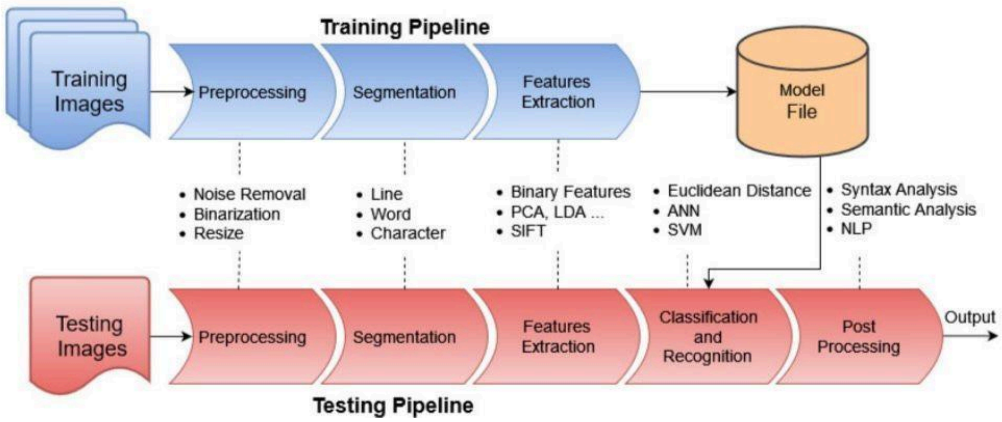


figure 2

These models can't replace human expertise for now, but they are valuable in providing hypotheses and in proof-checking human work in this field.

2. Limitations and challenges

Undeciphered languages stand as a classic example of a 'low-resource' case for NLPs, because they dispose of little to no datasets, no translated equivalents, and sometimes they are not related to any known languages. These limitations impede the application of standard traditional methods.

Some of the hindrances are the absence of spaces between words in ancient texts, the extremely reduced number of examples (only 7000 words for Linear A), physical degradation of manuscripts, and the uncertainty concerning phonetic interpretations and symbol meanings. Additionally, most undeciphered languages are not related to any known ones, which terminates the possibility of using parallel data. Modern NLP models usually require hundreds of thousands of examples, and in cases like these, only a few hundred exist. The evaluation of the results is also difficult, since we can not rely on a general truth about the meaning and the phonetics of the analyzed texts.

To overcome these issues, researchers use strategies such as data augmentation (the practice of artificially expanding training datasets through transformations or synthesis), knowledge transfer from related languages (where we can discuss related languages), and unsupervised alignment based on phonetic rules.

In addition to altering already existing images, researchers now use AI models to generate original synthetic handwriting samples. Using a technique known as GAN (Generative Adversarial Network), a model can learn the unique style from a small number of real samples and then produce other realistic text lines. When these artificial examples area added to the training data, they improve recognition performance for historical manuscripts that have very few original specimens available (Chang et al., 2021).

3. Results

Though the area is novel, NLP models have proven that they can accurately restore significant parts of the lost language, sometimes efficiently in terms of time, too, managing to finish their workload in a few hours. Hence, AI's potential for improving accuracy and efficiency is promising.

The Luo, Cao & Barzilay model deciphered 67% of Linear B vocabulary, proving that neural methods can learn the structure of an old language without parallel data. Coming to Ugaritic, the very same model has surpassed previous models with an absolute increase of 5.5% in establishing relations between known and extinct languages. Thus, AI models play a significant part in assessing hypotheses and proving or disproving theories about language family, phonetic, and morphological values. However, these results demand expert validation to identify and correct reasoning errors before scholars rely on them for broader tasks.

AI VS TRADITIONAL APPROACHES

1. Data dependency

Data dependency refers to the essential reliance of model performance on the scale, diversity, and quality of training datasets. Prior to the emergence of deep-learning models, researchers generated hand-crafted features useful in statistical methods and classification algorithms. A breakthrough in HTR research happened with the shift from feature-based models to transformer-based architectures, which allowed models to select and focus on the important segments of the analyzed texts (Sommerschield et al., 2023).

Low-resource constraints (sparse manuscripts, fragmented inscriptions, and limited parallel data across language pairs) fundamentally limit model generalization and prevent the supervised-learning approaches that dominate modern NLPs (Meoded, 2025). Additionally, reliable documents often suffer from real-world noise ensued by environmental preservation conditions, such as fading, corrosion, and smudging. Thus, there is a common scarcity and overuse of ideal textual manuscripts that could aid the training of the models. As manual transcription by humans has been most commonly used, it is a time and resource-consuming task. The importance of manual transcriptions must be emphasized, as they have assisted AI training in CNN models that achieved incredibly high accuracies. However, because it is a laborious and consuming job (AlKendi et al., 2024), researchers have resorted to augmentation (synthetic data generation) (Diao et al., 2025).

Neural networks often perform significantly better when they are trained on large and varied datasets, suggesting a structural dependency on these types of data. Taking these findings further, researchers have obtained comparatively better results when pairing these datasets with artificial ones (Sommerschield et al., 2023). This demonstrates the crucial role that augmentation plays in low-data settings, especially those simulating realistic manuscript degradation like Elastic Distortion (the warping of image pixels to mimic paper deformation). This can be useful, especially in the case of transformer models. Given their architectural complexity, they need more data diversity; they require a high volume of labeled data and extended training resources, which are tenuous to acquire.

Usually, datasets tend to be culturally narrow, because most historical documents preserved have religious or cult themes. Models adapt to the dominant style, and they lose their cross-domain generalization. Consequently, new directions appear, such as transfer learning (transferring knowledge learned from one task or language) and synthetic data generation. The augmentations in question work by simulating degradations in written texts, reducing CER (Character Error Rate - the percentage of characters incorrectly recognized) by half. Synthetic data can achieve near-real recognition accuracy, while embedded augmentation methods maximize learning in settings with low samples (Meoded, 2025).

2. *Interpretability*

There is no doubt that the availability of extensive datasets has aided the training and development of AI models in language deciphering by improving their accuracy. However, this introduced the issue of understanding how these models reason. The process of backtracking a model's decision process, known as interpretability, has hence become a central point in comparing AI approaches with traditional linguistic methods. Deep-learning models are usually referred to as 'black-box' models because they generate accurate predictions without exposing their decision logic, in contrast to traditional linguistic methods, where every step can be justified from a logical and philological stance. The recent development of these models has shown that as they become more powerful, the transparency decreases (Sommerschield et al., 2023).

Ground truth refers to reference translations that can be verified as real and viable, supplied by measurable data and analysis (Sommerschield et al., 2023). Without such a reference point, AI cannot be evaluated rationally. Errors can occur without the system being able to explain why, since the assessment of HTR models is demanding: errors can emerge from image noise, inconsistent character shapes and handwriting styles, and contextual misinterpretation (AlKendi et al., 2024). Without human interpretation, there cannot exist a precise way to evaluate outputs, especially since cultural materials require a personal and subjective judgement because no universal correct way of reading exists.

AI on its own is efficient, but opaque. Involving active human competence in the decision loop improves performance. HTR cannot stand on its own; it needs to be embedded in conventional human methods. The Ithaca project exemplifies this hybrid approach: the model achieved 62% accuracy in autonomous restoration of damaged Greek inscriptions, but when integrated into historians' workflows, allowing expert review and manual correction, accuracy reached 72%, demonstrating that human interpretation is essential (Sommerschield et al., 2023).

In the case of commercialized HTR models, such as Transkribus, there is a general lack of transparency. In order to embed HTR into practice, transparency, training, and datasets are the pillars of the process, as the human-AI collaboration entails some ethical responsibilities (only achievable via clear processes) in order to improve user trust when using automated systems. Critically, when AI is used for culturally relevant texts, it must not function in isolation from the human and historical context they emerged from (Terras, 2022).

3. Scalability

Scalability refers to the system's ability to adapt its complexity, size, and overall efficiency to accommodate a growing or decreasing amount of work. In AI text recognition, this issue appears in two cases: when limited resources that hinder the model's capacity to properly meet high requirements, or when it is needed to adapt an OCR model to another task (dealing with a different language or a different script) without training it from scratch.

Scaling these approaches to overcome the issues of scarce resources and multilingual domains is difficult because of the cost and time associated with creating transcribed data (Ingle et al., 2019). Additionally, increasing the capacity of these machines requires a generalization of their corpora, which is obstructed by the complexity of texts that vary geographically and chronologically, and across their genres. Hence, scalability is affected by low resources and the diversity of the reliable materials (Sommerschield et al., 2023).

Several solutions to improve scalability have been put forward. First, data augmentation is paired with improving the quality of existing datasets. Then, an architectural solution that aims to modify the way the model is built to handle more data better. Additionally, a cost-lowering alternative has also been tried, where researchers tried to lower the real data and focus on synthetic data (Ingle et al., 2019). This is currently the way in which many HTR models function (AlKendi et al., 2024).

Scalability is also a significant barrier when dealing with higly complex writing systems, such as arabic, which features cursive ligatures and characters with a meaning dependent on the context. In order to overcome these barriers while maintaining efficiency, recent development has moved towards specific transfomer architectures that optimize the extraction of features. For instance, Chan et al. (2024) introduced the HATFormer model, which uses self attention mechanisms to specifically map and model the structure of Arabic texts, which are based on the dependency of the characters. The architecture allows the model to properly work with many handwriting styles and diacritics. Hence, specifically designed transformers can overpass older recurrent models on complex scripts while saving up on computational power.

Recent technological developments have enabled some models to learn better and generalize their input and output range. However, the process is still constrained by concerns of time and cost (Ingle et al., 2019).

4. Cultural Context

Ancient languages represent a bridge towards ancient civilizations, as they embody ideas, cultures, and hidden histories. Therefore, understanding them requires a deep awareness of the context in which they appeared. Accurately assigning the place and time of ancient scripts is tied to allocating them their original and cultural context, a practice linguists have followed for decades, a process that needs to be embedded in machine learning as well. This can be particularly relevant given that the current representation of languages and histories is imbalanced, which has led to an event called 'digital colonialism'. An automatization of the decipherment process can bring an equilibrium to our current representations (Sommerschield et al., 2023).

ANALYSIS

1. Introductory framing

Artificial intelligence has transformed the study of ancient writing systems by expanding human capacities and digitalizing methods of reconstruction and visual interpretation. The emergence of deep learning enabled computers to identify linguistic and morphological patterns and, consequently, easily recognize cognates that would aid experts in the deciphering process. This digitalization revolutionized human-computer collaboration, as these technologies can complete and assist current methods used by linguists. The following cases - Linear B, Ugaritic, Etruscan, and Egyptian hieroglyphs - encapsulate this evolution from image-based recognition to interactive interpretation

2. Linear B

Linear B is a key case study, as it was the first syllabic system to be tested with CNNs, and partial prior knowledge regarding this language has aided machine learning.

Visual inputs for AI reconstruction have proven as particularly relevant in this case. As observed by Daggumati and Revesz (2018), who used a CNN and SVM to trace back Linear B's genealogy, Linear B is visually close to the Cretan Hieroglyphic script (Sommerschield et al., 2023). Moreover, data augmentation as a human-machine collaboration has been experimented with by adding the Linear B writing to existing data. Results suggest that Linear B may represent the key to solving decipherment issues, thanks to researchers' partial knowledge of it (Sommerschield et al., 2023), as long as it's used as assistive technology.

The importance of the dependency of human expertise cannot be undermined, particularly in neural decipherment. Architecture must be adapted to linguistic specifications, as a study by Luo, Cao, and Barzilay discovered. The researchers used a more general approach for decipherment, based on a sequence-to-sequence model, NeuroCipher (Sommerschield et al., 2023). Their input corpus consisted of the lost language (Linear B) and a non-parallel one (Greek). They manually modified the regularization to make up for the fact that two Greek letters correspond to one Linear B script. A misalignment resulted in a deletion error, as one of the letters was lost, since the sequence isn't perfectly aligned between the two languages. Their result is particularly relevant since Linear B is a syllabic language (Luo, Cao & Barzilay, 2019).

Thus, experiments involving Linear B, ranging from using it as an augmentation element and translating it using cognates, have shown that machine learning and result interpretation are reliant on philological skill. This dependence between machine inference and philological expertise stands as a foundation for subsequent work on Ugaritic.

3. Ugaritic

Ugaritic stands as a bridge between lost Semitic languages and a resource for linguistic models.

First attempts to use probabilistic translations were done by using a non-parametric Bayesian model, which worked by evaluating individual characters and morphemes in Ugaritic and how they correspond to certain cognates in Hebrew (Sommerschield et al., 2023). This model accurately simulated the reasoning of a linguist, within the computer's limits. This experiment set the base for other models that could learn how to automate the mapping process.

NeuroCipher, though more recognized for its achievement in Linear B translation, also achieved high performance results by capturing character correspondences between Ugaritic and Hebrew. It achieved a 5.5% improvement over previous models. Ugaritic is an alphabetical language, and Neurocipher's ability to decipher two structurally opposite languages represents a breakthrough, showing that the architecture can adapt and deep learn what connects languages across linguistic families (Luo, Cao & Barzilay, 2019).

These experiments have the potential to guide future learning directions; however, they are limited by the low-resource environments and the expenses of involving human experts.

4. Etruscan

Etruscan represents an extreme case of linguistic translation, where AI compensates for the scarcity of current data, as it is an ancient language spoken in Italy, from 700 BC to 50 AD, of which there are currently roughly 12.000 documented inscriptions.

In the Larth project, researchers have attempted to solve this matter at hand by adopting a three-step methodology. First, they decided to expand the existing datasets by gathering all the existing academic data and attempting to enlarge the input corpus through manual and automatic data augmentation. Then they experiment on Etruscan machine translation to English, assessing the possibility of training models using the artificially expanded datasets by comparing the performances of these models relative to others that lack diverse data. In the end, they assess the possibility of utilizing any similarities between Latin or Ancient Greek with Etruscan to upgrade the better-performing model.

They discovered that models worked better even with less data, suggesting that the transfer learning from Latin and Ancient Greek was effective. The models trained on Latin and Ancient Greek had more accurate results, having less difficulty in identifying lexical and morphological patterns. Although the resulting translation accuracy remains moderate, the models performed consistently better than random, confirming that meaningful linguistic connections can be found even in data-scarce conditions (Larth: Dataset and Machine Translation for Etruscan).

Consequently, they stabilized a reproducible methodology: the creation of corpora, neural training based on them, and the transfer of knowledge between ancient texts in various languages.

5. *Egyptian Hieroglyphs*

Egyptian Hieroglyphs represent one of the most complex writing systems in human history. They incorporate crucial information that could help society have a better understanding of historically and culturally relevant events, reasons for which this language has long been the focus of researchers and philologists (Diao et al., 2025).

The manual process brought along challenges in terms of time management and error frequency. Recent advancements in OCR algorithms based on deep learning promise to overcome these obstacles (Diao et al., 2025).

Early efforts in hieroglyphic recognition were based on image processing paired with a statistical language model. For instance, Franken et al. put forward such a system, utilizing a dataset composed of annotated glyphs from the Pyramid of Unas. Their result highlighted the importance of merging visual and linguistic models for hieroglyphic recognition.

Deep-learning algorithms have also been studied. Barucci et al. significantly improved the hieroglyph decipherment process by adopting such a model, a Glyphnet. This model incorporates and enhances several CNN architectures. Guidi et al. used a fine-tuned Mask R-CNN model, achieving highly accurate segmentations across varied visual sources. Thus, they created a base for transcription and stylistic analysis using a deep-learning method.

Plecher et al. developed ARsinoë, an augmentation system that combines an LSTM (Long Short-Term) memory with a 3D interactive design, to support recognition of both handwritten and printed hieroglyphs. Following this direction, Sobhy et. al proposed a complete interpretation framework, based on R-CNN detection, few-shot classification, and linguistic shaping. (Diao et al., 2025)

LIMITATIONS

Researchers currently have access to a very narrow number of ancient manuscripts due to preservation methods and time, and weather damage. Out of the surviving data, only a fragment has been digitized and adapted properly to the metadata format, which is a key factor in the deep-learning processes.

The machine learning approach and assessing the respective results are often obstructed by the lack of a verified reference dataset. Consequently, HTR models require a solid philological base and a very strong awareness of the cultural context.

The shortage of qualitative and quantitative data inhibits the implementation of recognition pipelines, which rely on the existence of well-grounded digital datasets. HTR thus becomes one of the most demanding challenges in ancient script restoration (Sommerschield et al., 2023).

The limitations go beyond the technicalities, and they can concern ethical and epistemic areas. At the moment, HTR technologies have reached a development level where they can be adopted by libraries, provided that they explain the thought process and results, allowing a safe implementation in historical procedures. However, the application of HTR is not regulated, and there has been limited research explaining the best practices involving the collection, distribution, and descriptions of the process behind the outputs.

Libraries that wish to generalize the use of HTR have the possibility to adhere to the procedures followed by technology companies, though they often lack transparency as well. For instance, Transkribus, an AI-driven platform for HTR, has significantly accelerated work; its architecture raises concerns regarding bias, as it does not fully qualify for the Open Science principles (Terras, 2022).

Recent studies on ancient text image recognition almost consistently emphasize that the physical conditions of the manuscripts are one of the most intense obstacles. Unlike modern text samples, ancient scripts suffer from time degradation (which results in images containing real-world noise) and from a scarce number of surviving scripts (which have shortened over time as a result of historical context, preservation environment, and the physical properties and behaviors of the materials used), resulting in a deficiency of reliable specimens. Additionally, character frequencies among these manuscripts tend to differ, as most texts treat a specific topic, usually relating to religious or historical events. This imbalance leads to an overfitting recognition, particularly when dealing with smaller datasets. Furthermore, image restoration and denoising algorithms operate at a pixel level, not being able to use true semantic understanding. This results in mixing up real strokes and noise that mimics them. Consequently, authentic strokes may be mistaken for real noise and be removed or, conversely, background noise to be mistaken for the text (Diao et al., 2025).

Papers in this domain have been exponentially appearing since 2014. This growing field demonstrates an academic interest largely fueled by the release of powerful, open-access HTR tools and the availability of large datasets. This trend suggests the research community is moving towards creating fully automated systems that are easier to use, verify, modify and even create (Swindall, Player, Keener, et al., 2022).

As research advances, addressing these limitations will entail both technical adaptation and a deeper integration of historical context and transparency. Collaboration between humans and new technologies will ensure that machines will enhance human interpretation.

DISCUSSIONS

Recent experiments confirm the value of these aforementioned techniques. Using data augmentations designed specifically for historical relevance, such as those that simulate faded ink or damaged paper, can reduce the CER by nearly half compared to a standard model. Researchers also found that merging outputs from different models improves the results even further. Consequently, using a diverse set of methods could potentially be another way to reach the highest accuracy results in historical HTR as efficiently as possible (Koch et al., 2023).

The analyzed studies have also shown a clear transition between manual decipherment and AI-assisted translation. Beyond accuracy concerns, the results raise questions regarding the epistemology of the machines, in regard to what the machines 'know' when they translate and reconstruct texts. Digital decipherment becomes a separate area of study, not just a tool, as researchers have demonstrated that it is still dependent on factors such as AI-human relationships that can safely interpret the outputs (Sommerschield et al., 2023). The results depend heavily on philological expertise that can accurately take into consideration subjective details such as the quality of data and the transparency of the results (Sommerschield et al., 2023).

The rapid shift between traditional and automatic interpretation is assigned to the digitization of the corpora data, having been made more accessible to the general public (Sommerschield et al., 2023). This allows a global partnership between researchers, which facilitates a more effective collaboration between human experts and machines. The collaboration is now essential for reproducing and improving the current models; it defines the future models and our current representation of them accepted as valid. Though the models are developed, scholars have to maneuver them in a very specific way to not leave room for interpretational doubts regarding the machine's reasoning process (Terras, 2022). Data transparency is as significant as the technology.

Regarding the transparency issue, AI models can generate highly accurate results that are not only unable to back up, but additionally unable to completely understand. This phenomenon was first investigated by Luo, Cao & Barzilay when they proved that even the most advanced models depend on pre-established rules, so even AI models learn relative to tradition. Their model's architecture was built on linguistic history patterns to make up for the limited supervision. It used a sequence-to-sequence algorithm to capture correspondences between cognates. Similarly, some Larth models successfully translated, but they provided no reasoning behind their process, nor did they guarantee any precision (Vico & Spanakis, 2023).

AI assistive models are not only technical tools, but they also strongly influence the way that knowledge is handled (produced, checked, and shared) among scholars. However, as technology continues to become commercialized, the processes might become completely inaccessible. This hinders the opportunity for reproducing and adhering to the principles of Open Science, which are fundamental for humanities-related research (Terras, 2022). Human expertise is the interpretive control guide of AI, and if companies fail to recognize this, the evolution of these systems might be damaged.

To attenuate the 'black-box' nature of the models and address transparency significance, scholars should shift their focus towards 'Trustworthy AI' frameworks, which are currently developing, especially in high-stake fields such as clinical medicine and engineering. Methodologies used in these domains offer guides for philology approaches. One valuable example would be the integration of XAI (explainable AI), specifically LIME (Ribeiro, Singh, & Guestrin, 2016), into transformer-based models in order to disclose the internal decision logic. Furthermore, another framework that aids scholars in explaining black-box features is the POMELO framework. It integrates a perturbation analysis technique (a method which allows researchers to explain outputs by identifying which input feature plays the most significant part in the model's translation, such as specific character sequences or partial images). Integrating these interpretability marks would transform AI from an opaque service to a transparent tool.

The future of AI in philology depends on transparency and interdisciplinary collaboration. Vico & Spanakis, 2023 argue that the publication of datasets could encourage further research on any language with low resources, aside from Etruscan. The future consists of an open collaboration and shared data networks. According to Sommerschield et al., machine learning can extract statistical patterns from large datasets, enabling new paths into the study of ancient scripts. AI is quantitative, but its power relies on the interpretability given by humans. It is up to scholars to give patterns a meaning.

CONCLUSION

This paper set out to assess the extent to which AI models, such as NLPs trained under low-resource conditions, can assist the deciphering lost languages, in contrast to traditional methods. The results reveal that while AI cannot be considered a reliable replacement for human expertise, it can considerably quicken the process of accurately restoring damaged scripts and point out important cognates. Across different studies focusing on ancient languages that were not provided large corpora of training data, AI models achieved measurable advancements. Technologies such as Pythia and NeuroDecipher highlighted the importance of a trained human eye in the digitization process for the best outcome. Widespread and commercialized models such as these raised concerns about transparency and issues of interpretability.

This study supports the idea that, despite several concerns, AI will not replace philologists in this area of study. Rather, it is a complementary tool that can reinterpret and complete our current ideas regarding ancient civilizations.

Future studies should focus on the potential of multilingual models, an area which has not been explored to its fullest, transparency, and ways to ethically reproduce the technologies.

REFERENCES

[1] Sommerschield, T., Assael, Y., Pavlopoulos, J., et al. (2023). *Machine Learning for Ancient Languages: A Survey*.

[2] Diao, Y., Xu, L., & Sun, W. (2025). *Ancient Script Image Recognition and Processing: A Review*.

[3] Chen, D., Shi, F., Agarwal, A., et al. (2024). *Comparing Visual and Textual Representations of Ancient Logographic Writing Systems for NLP*.

[4] Luo, J., Cao, R., & Barzilay, R. (2019). *Neural Decipherment via Minimum-Cost Flow: From Ugaritic to Linear B*.

[5] Tamburini, F. (2025). *Computational Approaches to Ancient Language Decipherment*.

[6] Coquenet, D., Chatelain, C., & Paquet, T. (2020). *Recurrence-free unconstrained handwritten text recognition using gated fully convolutional network*.

[7] Meoded, E. (2025). *Handwritten Text Recognition of Historical Manuscripts Using Transformer-Based Models*.

[8] AlKendi, W., Gechter, F., Heyberger, L., & Guyeux, C. (2024). *Advancements and Challenges in Handwritten Text Recognition: A Comprehensive Survey*.

[9] Hamdan, M., & Cheriet, M. (2023). *ResneSt-Transformer: Joint attention segmentation-free for end-to-end handwriting paragraph recognition model*.

[10] Chang, J.-H. R., Bresler, M., Chherawala, Y., Delaye, A., Deselaers, T., Dixon, R., & Tuzel, O. (2021). *Data incubation—Synthesizing missing data for handwriting recognition*.

[11] Terras, M. (2022). *The Role of the Library When Computers Can Read: Critically Adopting Handwritten Text Recognition (HTR) Technologies to Support Research*.

[12] Ingle, R. R., Fujii, Y., Deselaers, T., Baccash, J., & Popat, A. C. (2019). *A Scalable Handwritten Text Recognition System*.

[13] Han Chan, A., Mijar, A., Saeed, M., Wong, C.–W., & Khater, A. (2024). *HATFormer: Historic Handwritten Arabic Text Line Recognition with Transformers*.

[14] Daggumati, S., & Révész, P. (2018). *A Statistical Model for the Decipherment of Minoan Linear A*.

[15] Vico, F., & Spanakis, G. (2023). *Larth: Dataset and Machine Translation for Etruscan*.

[16] Sobhy, M., El-Sayed, S., & Hussein, H. (2023). *Advancements in Handwritten Text Recognition for Historical Manuscripts*.

[17] Swindall, M. I., Player, T., Keener, B., Williams, A. C., Brusuelas, J. H., Nicolardi, F., D'Angelo, M., Vergara, C., McOsker, M., & Wallin, J. F. (2022). *Dataset augmentation in papyrology with generative models: A study of synthetic ancient Greek character images*.

[18] Koch, P., Nuñez, G. V., Arias, E. G., Heumann, C., Schöffel, M., Häberlin, A., & Aßenmacher, M. (2023). *A Tailored Handwritten-Text-Recognition System for Medieval Latin*. In Proceedings of the Ancient Language Processing Workshop.

[19] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). *"Why Should I Trust You?" Explaining the Predictions of Any Classifier*.

[20] Coquenet, D., Chatelain, C., & Paquet, T. (2021). *DAN: a Segmentation-free Document Attention Network for Handwritten Document Recognition*.

[21] Plateau–Holleville, C., Bonnot, E., Gechter, F., & Heyberger, L. *French vital records data gathering and analysis through image processing and machine learning algorithms*.