EDICIONES Universidad de Valladolid

AI Applications in Hellenic Studies. A Survey*.

Aplicaciones de IA en Estudios Helénicos. Una revisión.

DIEGO CHAPINAL-HERAS
C/ Francisco Morales Nieva 1, Madrid. 28049. Módulo II, 3.10
diego.chapinal@uam.es
ORCID: 0000-0002-6992-184X

CARLOS DÍAZ-SÁNCHEZ
C/ Profesor Aranguren s/n, Madrid. 28040
cardia1294@gmail.com
ORCID: 0000-0001-5706-8733

Recibido: 14/06/2024 Aceptado: 30/09/2024
Cómo citar: Chapinal-Heras, Diego y Díaz-Sánchez, Carlos, "AI Applications in Hellenic Studies. A survey", *Tabanque. Revista pedagógica, 36* (2024): 49-71.
DOI: https://doi.org/10.24197/trp.36.2024.49-71

Resumen: Este artículo reúne los avances más importantes en Inteligencia Artificial (IA) en el ámbito de los estudios helénicos, centrándose especialmente en la Literatura, Paleografía, Epigrafía, Arqueología e Historia del Arte. Se analiza el surgimiento de varias aplicaciones, junto con el software y las metodologías empleadas. El principal objetivo es presentar los desarrollos más notables en la investigación, al mismo tiempo que se evidencia el limitado número de casos de estudios disponibles. La integración de la IA tiene un gran potencial que, paulatinamente, se va reconociendo en este campo. Esta revisión del estado actual de la investigación pretende demostrar de qué manera nuestro ámbito de trabajo puede beneficiarse de los diferentes enfoques adoptados, junto con su potencial para futuros avances. en la edición de un corpus de inscripciones griegas que presenta un alto nivel de complejidad.

Palabras clave: Inteligencia Artificial, tecnología con IA, estudios helénicos, Ciencias Humanas, Humanidades Digitales, Historia, Paleografía, Epigrafía, Arqueología, Historia del Arte.

Abstract: This paper brings together the most significant impacts of artificial intelligence (AI) technology in the sphere of Hellenic studies, particularly focusing on areas such as Literature, Palaeography, Epigraphy, Archaeology and Art History. It discusses the emergence of various applications along with the software and methodologies used. The primary aim is to showcase the key advancements in research while also pointing out the limited number of case studies available. AI integration presents numerous opportunities that are gradually being embraced. This comprehensive overview of the current state of affairs aims to demonstrate how our field of expertise can leverage various approaches undertaken, along with their potential for future progress.

*This research has partly been carried out within the framework of project RYC2021-031612-I, derived from a Ramón y Cajal postdoctoral contract.

Keywords: Artificial Intelligence, AI technology, Hellenic studies, Human Sciences, Digital Humanities, History, Palaeography, Epigraphy, Archaeology, History of Art.

Summary: 1. Introducción. 2. Metodología. 3. IA en los estudios Helenísticos. 4. Discusión y Consluciones.

1. INTRODUCTION

The aim of this paper is to review the literature on Artificial Intelligence (AI) in the field of Hellenistic studies, focusing on a number of disciplines including Literature, Palaeography, Epigraphy, Archaeology and Art History. AI has been further applied in other areas of knowledge for a variety of purposes, such as developing virtual scenarios (Kolve et al., 2022), improving business operations (Makridakis, 2017; Jarrahi, 2018) and healthcare (Vaishya et al., 2020). Education is another sphere in which artificial intelligence (AI) is relevant, corroborated by the fact that UNESCO incorporated it as a component of the 2030 Agenda for Sustainable Development (UNESCO, 2019). Nevertheless, in the field of the Humanities, the advancement has not progressed at the same pace as in other fields.

The purpose is to provide a survey with descriptions of AI applications in Hellenic studies that are understandable for readers unfamiliar with Computer Science and Machine Learning (ML), enabling them to grasp the value and prospects of such endeavours as much as possible.

The significance, functions and applications of AI are diverse. Primarily, its objective is to empower machines to accurately interpret external data, to learn from such interpretations autonomously and ultimately to apply the acquired knowledge to specific tasks (García-Serrano and Menta Garuz, 2022). This process involves the collection, categorisation and management of extensive datasets. By using various algorithms, not only can computers execute tasks with precision but also engage with agents that use this resource or tool and adapt themselves to enhance efficiency (Bengio 2009). For a more comprehensive exploration of AI's potential and opportunities, readers are directed to the succinct article authored by LeCun, Bengio and Hinton (2015).

The progress of AI has traversed multiple stages, progressively integrating additional resources and broadening its capabilities. As will be demonstrated, this trajectory has directly influenced its use in the sphere of the Human Sciences, particularly within the domain of Hellenic studies. For instance, originating from text processing to identify terms and transliteration

capabilities, we have progressed towards the potential restoration of fragmentary inscriptions, either partially or in their entirety. Similarly, stemming from the objective of workload reduction, AI now demonstrates a remarkable ability to rival the proficiency of professionals across diverse disciplines, although in general an expert is required to review the work performed.

This pivotal advancement has primarily been propelled by significant strides in Deep Learning (DL), which seeks to empower machines to emulate human brain-like thinking and learning processes. For this objective, researchers use neural network models capable of processing data simultaneously across various levels (Bengio, LeCun and Hinton, 2021). In this domain, Convolutional Neural Networks (CNNs) stand out as particularly beneficial for virtual analysis and image processing. CNN models are applicable in both supervised and unsupervised learning settings. In supervised learning, the system is provided with input data and corresponding desired outputs (true labels), enabling the model to learn the mapping between them. Conversely, in unsupervised learning, the true labels for a given set of inputs are unknown, prompting the model to estimate the underlying distribution of the input data samples (Khan et al., 2022). Another significant advancement in this sphere is Natural Language Processing (NLP), which encompasses tasks such as text restoration. NLP involves the creation of computational mechanisms that enhance communication between humans and machines via language, leveraging algorithms and diverse methods for automatic ML (Pagé-Perron et al., 2017).

It is crucial to emphasise that the implementation of Artificial Intelligence (AI) technology in the Humanities does not imply replacing human labour. Rather, all the studies examined in this paper concur that AI's primary objective is to function as a tool that streamlines and enhances research processes. Computers enable this with their ability to process vast amounts of data much more quickly than human experts in any given field. Consequently, this progress should be viewed as a valuable asset in Humanities research, instead of perceiving it as a hindrance or potential threat to our work. In fact, in many instances, it is imperative to have a specialist to review AI-generated outputs to identify and rectify any potential errors, as achieving 100% accuracy in procedures involving AI is uncommon. This review process entails a less time-intensive task, bypassing the necessity to begin the entire process over again. The utility and advantages of CNN and NLP, among other methodologies, are indisputable, particularly in fields requiring the processing of substantial data volumes.

The use of these new methodologies and resources is essential for all areas of knowledge. According to Makridakis' (2017) perspective, during the course of this generation, the advance in AI will be so significant that we will be facing a technological revolution even more transcendental than the Industrial Revolution. With this scenario in mind, we need to encourage this development in the Humanities in general, and in Hellenic studies in particular. In recent decades, the progress of digital technology has brought about a substantial transformation in many areas of our field of study. It is however a process that is far from over. The literature cited below illustrates various ways in which we can implement AI approaches in different research areas in Humanities (Lee and Kim, 2020; Mantovan and Nanni, 2020; García-Serrano and Menta Garuz, 2022; Chapinal-Heras & Díaz-Sánchez, 2024; Díaz-Sánchez & Chapinal-Heras, 2023). The combination of different methods and techniques has yielded valuable results. Without these technological advancements and relying solely on traditional methodology, obtaining such results would have taken several decades.

2. METHODOLOGY

Two primary methods were used in the search for relevant publications regarding the topics under investigation in this study. The first port of call was bibliographic repositories, with Semantic Scholar and Google Scholar being the main platforms. Specific search queries were entered in these repositories, including "AI Hellenic studies," "Deep Learning Hellenic studies," "AI Greek Palaeography," "AI Hellenic Epigraphy" and "AI Greek Archaeology". The second method involved the compilation of additional bibliography sourced from the most recent publications identified previously. Through this approach, we have assembled the contributions developed during the 21st century in the field of Hellenic studies.

3. AI IN HELLENISTIC STUDIES

Chronologically, the survey covers mainly the Bronze Age to Late Antiquity; and geographically, the core territory of the Greek civilisation, i.e. the

Helladic Peninsula, the Aegean area and Asia Minor[1]. A few exceptions belong to topics where documents have a later date or archaeological data is from more distant regions or has an earlier dating but specific connections with Greek culture and its formation. The organisation of the literature below is based on wide-ranging study areas, more specifically Literature, Palaeography, Epigraphy, Archaeology and Art History. Each section describes, in chronological order of publication, the different applications that have been developed. In the sphere of Literature and Palaeography, despite the focus being placed on the Greek language, the survey included specific research applied in Classical Latin, as both are closely connected and most Latin authors dealt with aspects related to Greek culture.

3. 1. Literature

In Greek and Latin Literature, the implementation of the Perseus project at the end of the 20th century (Smith et al., 2000) proved the opportunities offered by computer technology for working with texts and their various associated components. However, the application of AI has come about more recently. The launch of BERT (Bidirectional Encoder Representation from Transformer) by Google in 2018 opened up new perspectives for classical language studies. The main function of this technique is the pre-training of NLP to enable the system to improve its ability to analyse and interpret user queries. The key feature of BERT, which marked a breakthrough in NLP, is its contextual embedding model. This means that the analysis of each word takes into consideration the other words in the sentence. This provides a more accurate understanding, especially in the case of words with multiple meanings, depending on the context (Devlin et al., 2019; Ravichandiran, 2021). The innovation of BERT in the field contextual language model in English has led to the development of specific applications in other languages. The most significant case in terms of this article is Latin-BERT created by Bamman and Burns (2020), which uses Latin sources spanning from 200 BC to the present day. Classical Latin is no longer spoken, but it has been used and has evolved in different contexts over this long period. This has allowed BERT to be trained with 642.7 million tokens. With this bulk of data, the software can correct fragmentary texts by providing word probability

[1] Although the scope included the Greek colonies distributed throughout the Mediterranean and the Black Sea, we have not found any specific research with AI that examines these settlements.

estimates, allowing researchers to identify the most likely and accurate option in each case.

One of the main branches of AI research in Literature is Topic Modelling, which consists of identifying the most relevant topics associated with a collection of documents. This unsupervised mathematical model takes as input a set of documents D and returns a set of topics T that accurately and consistently represent the content of D (Churchill and Singh, 2022). A noteworthy application is that of Koentges (2020), who set out to identify recurrent conceptual elements in the philosophical schools of thought present in the first thousand years of ancient Greek literature. To carry out this research, Koentges used the LDA Topic Modelling model, which allowed him to generate documentary vectors based on the structural and thematic content of each text analysed. The researcher used the open access software (Meletē) Tōpan, which he created himself, to perform this study. The document collection examined for this purpose contained almost 30 million words. It is important to bear in mind that the language processed is Ancient Greek, not a current and spoken language. This posed an additional challenge for the analysis. However, the results obtained were highly positive, demonstrating that with AI it is possible to identify references that include aspects of philosophical thought from that era.

Another recent step is the Logion Project, at Princeton University and conducted by Cowen-Breen et al. (2023). The initiative is to develop an NLP tool that aids the restoration and elucidation of premodern Greek texts. It started with the Byzantine author Michael Psellos, and is currently expanding to other Greek authors such as Aristotle or Galen[2].

3. 2. Palaeography

As in Literature, the field of Palaeography has benefited from the application of AI in the identification of images, symbols, signs and characters to develop its analysis. Following research begun a few years ago (Arabadjis et al., 2013), Arabadjis, Papaodysseus and Mamatsis (2021) developed a program capable of processing, grouping and comparing 2D shapes with shapes of the same class. It facilitates the identification of authors according to the type of writing and the individual characteristics observed in the different letters. Using DL, this team has been able to cluster documents according to their writer, determining the maximum clusters that also maximise the joint probability of

[2] https://logionproject.squarespace.com/ [consulted 28 May 2024].

classification calculated on the characters that appear in all documents. The processing they developed focuses on the visible deformations within the palaeography of these documents: namely, each deformation is divided i) into a component that affects the congruence class of a given shape as a whole, and ii) into a component that represents the deformations of a shape within its class. By neutralising the deformations of the second class we obtain an overall mutual alignment of the shapes of a given group. Then, the deformations of the first class represent the deviations of the shapes within the group and their absorption results in the determination of the representative shape of the group. To test of this system, the methodology was applied to the automatic identification of authorship in medieval manuscripts, more specifically 26 images of Byzantine manuscript pages that contain sections of the Homeric Iliad. They were selected by Prof. Ch. Blackwell of Furman University in order to test the effectiveness of the system. The results of this analysis concluded that there were four different writers. In order to check that their conclusions were reliable, Prof. Blackwell evaluated them and noted that AI had established the same classification as previous scholars, except in one case, where the system had linked two manuscripts as belonging to the same writer. This last identification, explained Prof. Blackwell, confirmed an existing hypothesis, not corroborated until now (Arabadjis et al., 2021: 183).

New developments implementing AI with epigraphy are aimed at the automated identification of characters and legible forms to analyse, recognise, recover and restore writings. Griffin (2023) uses the term "artificial palaeography" to refer to the application of ML technologies for Optical Character Recognition (OCR) and Handwritten Text Recognition (HTR) in ancient written works. Automated transcription of ancient manuscripts would free scholars from the tedious tasks that must be performed before analysis and interpretation can begin, providing them with the means to search, retrieve and access content in the same way as in a modern text archive. Examples of this type of methodology include the Research Environment for Ancient Documents (READ) platform, an open-source web platform that offers a number of tools for converting images of orthographic units for transcription; and the EU-funded Transkribus Project, which involves the application of AI to convert images of many handwritten texts, including medieval Latin texts, into machine-readable text with a relatively low error rate.

The use of AI for text reconstruction as a non-invasive method of text reading can be seen in the experiments carried out by Parker et al. (2019). They

developed a novel method to recover and improve the ink of the Herculaneum papyri. Charred by the eruption of Vesuvius in 79 AD, the papyri are fragile and difficult to read due to ink degradation. The authors presented a non-destructive technology that uses X-ray phase contrast tomography (XPCT) and ML to reveal the invisible ink and increase the readability of the text. XPCT captures high-resolution 3D images of papyri and ML algorithms recognise and segment ink traces from the background to improve visibility.

3. 3. Epigraphy

In the last two decades, Epigraphy has seen significant advances in computer-based analysis, as the survey of Sommerschield et al. (2023) shows. Tracy et al. (2007) focused on identifying the hands of the letter engravers in a collection of inscriptions from ancient Athens. Although this first study was based on only six documents, it demonstrated the potential of technology in converting letter strokes into mathematical formulae. Technically this approach is not considered an AI application, but it involves the development of mathematical operations to process these types of sources, representing an important step towards the use of more advanced techniques with ML.

Subsequently, a team composed of some of the same authors expanded and improved this methodology. In a study of 24 Athenian inscriptions, they managed to identify six different writers (Panagopoulos et al., 2009). In a later study, 32 new epigraphs were tested by Tracy to enable analysis of the identity of the people mentioned in the inscriptions, the craftsman or the place where these materials were found (Papaodysseus et al., 2010). The results from the AI analysis facilitated the identification of up to nine different craftsmen who produced these epigraphs (Rousopoulos et al., 2011). It is worth pointing out that the methodology proposed in this research offers very promising results, but it must be borne in mind that the letters do not convey the idiosyncrasies of each writer in an equivalent way.

Amato et al. (2016) investigated Greek and Latin inscriptions within the framework of the EAGLE project, which collected a large number of epigraphs from both civilisations. The team focused on Fisher Vectors (FVs) to encode local descriptors, as well as on various CNN representations, including a combination of both approaches. Using these methods, they performed a visual similarity search among all images in the dataset to enable the system to recognise objects in query images. Following the conclusion of the EAGLE project, the International Digital Epigraphy Association (IDEA) was established for the purpose of maintaining the EAGLE resources and

continuing to promote cooperation and the integration of resources in the field. Its objectives also include overcoming the limitations of individual projects and moving towards the creation of an epigraphy info resource based on the model used by papyrologists. IDEA intends to continue the networking efforts of the EAGLE project and support its results, focusing on keeping the EAGLE portal infrastructure and its functionalities up and running, providing support to members who wish to contribute, advising new projects on the availability of resources, keeping abreast of developments in the field, and sharing this knowledge to promote more efficient and organised work in the field of digital epigraphy (Liuzzo, 2019).

Luo, Cao and Barzilay (2019) used a neural decryption algorithm to identify cognates in Near Eastern Ugarit texts and in the Mycenaean Greek Linear B. One of the main advantages of this application is its potential to develop an automated decipherment system that would facilitate the work of epigraphers by allowing them to work faster and with less effort, thus in the line of the aim of AI applications. The authors point out that, following the proposed systematic methodology, this approach could be replicated for any other ancient script with minimal adjustments.

In Greek Epigraphy, the variability of scripts, languages and writing styles and insufficient or missing information that requires educated guesses or extrapolations to fill in the gaps can make decoding and reconstructing difficult. DNNs have revolutionised the field of ancient text restoration by offering methods that are faster, more accurate and less labour-intensive than traditional approaches (Ali, Baheeja and Asia, 2023).

In this sense, the field has made significant recent advances thanks to the group led by Assael, Sommerschield and other colleagues. They first developed PYTHIA, a fully automated DL model trained to restore the text of ancient Greek inscriptions by predicting the character sequences that make up the hypothetical restorations. This was complemented by PHI-ML, a machine-processable text dataset consisting of over 3.2 million words (Assael, Sommerschield and Prag, 2019). This research has been extended and strengthened by the development of Ithaca, a DNN model that automates restoration and attribution, aiming at improving the analysis and restoration of epigraphic documents written in ancient Greek. Like PYTHIA and other previous studies in this field, Ithaca is based on the existence of formulas, patterns and expressions that tend to repeat. Using textual and contextual parallelism, the system can identify the geographical and temporal location of these inscriptions with a high degree of accuracy. This work was made possible by the creation of a dataset of machinable epigraphic texts containing

76,608 texts. Once the results have been obtained, the task of the researchers is to verify and check whether the reconstructions proposed by the program are sound or whether modifications are necessary (Assael et al., 2022)[3].

Another example of the use of these technologies and the potential of OCR in epigraphic studies can be seen in the GLAUx Project ("the Greek Language Automated"), an ongoing effort to develop a large, long-term diachronic corpus of Greek spanning sixteen centuries of literary and non-literary material annotated with NLP methods (Keersmaekers, 2021). However, this proposal to create an automated epigraphic corpus is still in progress, although data relating to its quantification has allowed the identification of patterns in the writing of some Greek epigraphs.

A recent innovation in Greek Epigraphy is the NextGenerationEu Project conducted by Martín González, "Dodona viaja a Ítaca. Inteligencia Artificial aplicada a la edición de inscripciones griegas" ("Dodona travels to Ithaca. Artificial Intelligence applied to the edition of Greek inscriptions)"[4]. Currently under development (2024-2026), this initiative is focused on the tablets of Dodona, a Greek oracle located in Epirus, where those requesting counsel wrote their questions on small, thin lead plaques. By means of the DNNs developed in Ithaca, the aim is to apply this technology to the critical edition of the oracular tablets carried out by the Dodona Online Project in recent years[5], therefore providing a hybrid analysis that combines the work of expert scholars with that of Ithaca. The expected results will be more accurate reconstructions and data concerning the dating of the inscriptions and the possible origins of those consulting the oracle.

Another ongoing initiative is the LACUNAE Project, conducted by Magnani and Zunino, at Università Degli Studi di Udine. It is dedicated to the creation of a digital tool that exploits AI in order to develop suggestions for the integration of gaps, lacunae, in Greek and Latin epigraphic documents based on a comparison with selected epigraphic documents and literary texts.

[3] The processed dataset I.PHI is a pipeline for downloading and processing the Packard Humanities Institute's database of ancient Greek inscriptions, including the geographical and chronological metadata, into a machine-actionable format: https://github.com/sommerschield/iphi [consulted 10 May 2024].

[4] https://clasicas.uva.es/docentes/elena-martin-gonzalez/ [consulted 22 May 2024].

[5] https://dodonaonline.com/ [consulted 22 May 2024].

Precisely for Greek inscriptions it aims to adapt and apply, strengthening it, the model already developed by the Ithaca project.[6]

3. 4. Archaeology

Similarly to that observed in the other disciplines, AI application in Archaeology is relatively recent and focuses on the analysis of databases – GIS, archaeological data– through ANNs. Databases have usually been constructed following the same schemes and have tended to compile imprecise and heterogeneous data. Computer Sciences have contributed to a better classification and organisation of this data. The following step, ANNs implementation, has enhanced the processing of such an enormous amount of information collected in datasets, using more complex analyses in which interconnectivity is the key. ANNs are based on an abstract analogy with the architecture and functionality of the brain. Brain synapses are branched connections specialising in electrochemical transmission (or outputs). On the other side of the neuronal cell, dendrites represent the section specialised in receiving signals (or inputs). It is therefore clear that neurons are cells specialised in reciprocal interconnection. The key factor for the development and application of ANNs, as well as for the brain, is the above-mentioned interconnectivity. The extent of interconnectivity determines the ability of the ANN to adapt to learning patterns. Due to their structure, ANNs are based on a fuzzy logic paradigm, which implies that they can become a very flexible tool, something that traditional Boolean-based algorithms would never be able to do (Deravignone and Jánica, 2006).

Recent years have seen a remarkable increase in studies in which telematic prospecting and remote sensing with AI has proved fruitful (Sharafi et al., 2016; Soroush et al., 2020; Davis et al., 2021; Argyrou and Agapiou, 2022; Küçükdemirci and Sarris, 2020). More specifically, predicting the location of archaeological sites has enormous potential. Argyrou and Agapiou (2022) outline its importance in taking decisions regarding appropriate conservation and protection strategies in terms of where best to excavate in a complex cultural landscape. On the other hand, as Jamil et al. (2022) point out, DL consumes large amounts of data to achieve greater accuracy and there are not many large datasets available for the field of archaeology.

[6] https://dium.uniud.it/it/ricerca/progetti-corso/digital-humanities-e-intelligenza-artificiale/lacune-epigrafiche-greche-e-latine-strumento-di-suggerimento-delle-integrazioni/ [consulted 28 May 2024].

In the region of Thessaly, the processing of satellite and archive data and images predicted the locations where Neolithic settlements could be detected (Alexakis et al., 2011). In another project, the team of Balla et al. (2012a; 2012b; 2013; cf. Pavlidis, Ben-Yosef and Jones, 2023) used computational image analysis through DL and ANNs in order to develop predictive models in the Macedonian region. The model processed archaeological and geospatial information from extensive literature research on the specific case of ancient Macedonian tombs in northern Greece and was able to provide map-based predictions with colour-coded probability visualisations. The model used multi-dimensional data and parameter selection and tuning to relax the imposed criteria and provide conservative or ambitious predictions. The fine tuning of parameters was extensively tested on real field data and several configurations were proposed for different application scenarios.

Orengo and García-Molsosa (2019) introduced ML in unmanned aerial vehicles that use high-definition remote sensing systems to survey areas more quickly and detect features on the surface or just below it. They applied this research in the Archaeological Project at Abdera and Xanthi (APAX) in north-eastern Greece, aiming to develop an automated pottery sherd recording methodology. The steps of the workflow consisted of the integration of drone high-resolution photography to record in detail the surveyed fields, performing photogrammetry to join all these photographs in a single orthophotomosaic. ML and other geospatial analyses were applied to identify and isolate pottery sherds in the photomosaic and cloud computing was used to accurately process these analyses.

Küçükdemirci and Sarris (2020) undertook one of the first attempts to generate a large, labelled dataset and trained network from scratch using a GPR depth slice data set archive. Their project used a convolutional neural network (CNN) built with the Python 3.6 programming language using the Keras DL Library with Tensorflow backends, a library that implements the building blocks for CNNs. The network was trained from scratch, adopting U-Net architecture to accomplish an automatic analysis of the archaeo-geophysical features with emphasis on ground-penetrating radar (GPR) anomalies. They focused on a selection of 800 images from different areas of mainland Greece, primarily Lechaion, Heraia and Mantinea in the Peloponnese (southern Greece), Halos in Thessaly (central Greece) and Olynthos in Chalkidiki (northern Greece).

Another example of the computerisation of archaeological data by means of AI has been the typological study of the pottery of Cyprus in ancient times, where the team of Agapiou, Vionis, and Papantoniou (2021) examined the

typologies of a single site through a visual analysis, shortening the time taken to study the pottery. In the same line of research, Hatzinikolaou et al. (2023) used ANNs and a fuzzy logic model to perform a spatial analysis on the island of Melos to test the effects of analysis with AI on the surveys carried out between 1974 and 1977 by the British School of Archaeology. The positive effect of this study made it possible to propose possible locations of sites in the territory based on the analysis of the maps and models developed in the 1970s. The study also showed that this method is an alternative that could contribute significantly to archaeological research that needs such data, especially in countries with a rich historical past, as it allows sites to be selected for surveying with a greater statistical basis and greater speed.

The use of AI is not only relevant for archaeological landscape studies or predictive modelling of settlements or burials, but also for underwater detection, as the work of Paraskevas et al. (2023) reflects. This team has created a ML model designed to identify ancient pottery sherds near a submerged shipwreck off Modi Island in Greece. Their primary goal has been to integrate the resulting object detection system into a remotely operated vehicle (ROV) for automated pottery sherd recognition, thereby aiding archaeological excavations.

Another area in which it is possible to benefit from AI applications is online sources. Iyer and Franklin (2022) have used ANNs and DL to train an AI capable of identifying sets of artefacts and developing questions or issues to facilitate access to this information, testing the methodology with the British Museum's online collection, including finds from the Greek culture. This AI application for specific web scraping has been further developed by Demir, Boyoğlu and Kayikci (2023), who, as part of the preliminary phase of an archaeological research project, proposed a workflow that can determine any architectural or other related ancient element more quickly than the traditional manual selection method. For example, the presence of a majority of columns in an ancient town indicates that it was a settlement with certain degree of development, as it shows the presence of numerous monumental structures with columns. For the analysis they used Selenium, a free, easy-to-use online tool, a framework that includes several resources and libraries that allows the automation of web browsers. It can work with any web browser without human interaction. A part of the Python code allows the user to search for information in different file types, such as text, images or any other multimedia format. The other main tool used in this study is Amazon Rekognition. Rekognition Custom Labels is Amazon's system of tens of millions of trained images in many categories. The content of the photos

uploaded to the system is automatically identified using integrated cloud-based deep learning. Amazon Rekognition is a visual analytics service that was launched in 2016 and has been used in a variety of fields since its inception. Providing two options for use, it offers algorithms pre-trained on data owned by Amazon and an algorithm base that the user can create by training it on their own data. As a result, a researcher studying ancient towns and columns can determine which of those in the region under study was a metropolis from the number of columns and obtain preliminary information about the styles of ancient columns in this region. A minor handicap is that the preparation phase requires the labelling of tens of thousands of images to run the DL model accurately.

In another vein, Maravelia et al. (2019), in the framework of the Hellenic Institute of Egyptology (HIE) Mummies Project, propose the use of DL by means of CT-Scanning imaging in conjunction with Computed Tomography (CT) for the study of Ptolemaic mummies in the National Archaeological Museum of Athens. The application of Smart Computing and its learning methods has allowed the identification of pathologies and the reconstruction of the anatomy of the individuals. The team has studied the age and sex of the mummies and compared them with the inscriptions found on their sarcophagi. It has also examined the methods of mummification and the concomitant alterations related to the preparation of the bodies for burial and revealed the comparative damage suffered by these mummies post mortem during the centuries since their death. Likewise, they have discovered the diseases and syndromes suffered by those individuals, as well as information related to their dietary habits.

3. 5. History of Art

The focus in this case has been the identification and study of iconography based on image analysis. Unfortunately, this type of research is scarce in Hellenic studies. A recent project undertaken by Wayenberg and Capriotti (2014) used ANNs to examine the figures of maenads and satyrs on red-figure pottery, focusing on their multiple interactions. The project was applied to a dataset of 114 vases with 96 variables related both to form (period, painter, function) and figures (appearance, hairstyle, behaviour, kinship). The methodology is based on the inclusion criterion of the presence of at least one satyr and one maenad identified on each vase. The researchers first selected the vases in the Beazley Archive following a stratified statistical method by periods (10% of all vases from each period), based on the postulate that a 10%

sample should be considered statistically relevant. They then organised a database with a total of 478 figures of maenads and satyrs. This initiative demonstrates the extent to which AI applications in iconographic studies facilitate the analysis of artistic expression in order to provide innovative and important interpretative solutions.

4. Discussion and Conclusions

This article has attempted to offer an updated compilation of AI applications in Hellenic studies in the fields of Literature, Palaeography, Epigraphy, Archaeology and History of Art. It is clear that their adoption remains limited compared to the trend in other areas of knowledge, a trend that can be seen in general terms in the Human Sciences. Several factors explain this situation: for instance the cost, the complexity of the processes and algorithms involved in programming AI in fields where Computer Science is usually considered secondary, and the scarcity of open-access datasets. On the other hand, the very existence of a number of projects and studies that show the potential of this type of research, together with the fact that AI has seen a remarkable growth in popularity in our society in recent years, is enough evidence to assume that in the short-term AI applications in Hellenic studies will increase. A good example is the TALOS project, a multibranch initiative funded by the European Commission (ERA Chairs HORIZON-WIDERA-2022-TALENTS-01) and developed at the University of Crete between 2023 and 2028. Led by Prof. Ch. Roche, the project focuses on AI and Digital Humanities, aiming to provide an optimum environment for both training and research for students, the academic community and society at large.[7]

We have described several case studies in which experimentation with different methods and techniques developed in recent years has contributed significantly to the advancement of data analysis. Tasks that usually require several years for an expert can be carried out by trained AI software in a matter of weeks or even days. The role of AI as a tool intended to assist the researcher, rather than to replace him or her, outlines the usefulness of such approaches. AI is not a technology that can automatically examine a dataset without human intervention at any step of the process. Scholars are essential to evaluate the results produced by the software in order to modify, supplement or extend them. Moreover, it encourages interdisciplinary collaboration, especially with the field of Computer Sciences.

[7] https://talos-ai4ssh.uoc.gr/ [consulted 22 May 2024].

BIBLIOGRAPHY

Agapiou, A., Vionis, A. and Papantoniou, G. (2021), "Detection of archaeological surface ceramics using deep learning image-based methods and very high-resolution UAV imageries", Land, 10, p. 1365. https://doi.org/10.3390/land10121365

Ali, A.A., Baheeja, K. and Asia, M.N.A. (2023), "Ancient textual restoration using Deep Neural Networks: A literature review", in *Al-Sadiq International Conference on Communication and Information Technology-2023 (AICCIT-2023)*, Baghdad, Al-Sadiq University, pp. 64-69. https://doi.org/10.1109/AICCIT57614.2023.10218159

Amato, G., Falchi, F. and Vadicamo, L. (2016), "Visual recognition of ancient inscriptions using convolutional neural network and fisher vector", *JOCCH,* 9(4), pp. 1-24. https://doi.org/10.1145/2964911.

Arabadjis, D, Giannopoulos, F., Papaodysseus, C., Zannos, S., Rousopoulos, P. Panagopoulos, M. and Blackwell, C. (2013), "New mathematical and algorithmic schemes for pattern classification with application to the identification of writers of important ancient documents", *Pattern Recog.*, 46(8), pp. 2278-2296. https://doi.org/10.1016/j.patcog.2013.01.019

Arabadjis. D., Papaodysseus, C. and Mamatsis, A.R. (2021), "Identification of the writer of historical documents via geometric modeling of the handwriting", in *2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Kuala Terengganu, IEEE, pp. 179-184. https://doi.org/10.1109/ICSIPA52582.2021.9576800

Argyrou, A. and Agapiou, A. (2022), "A Review of Artificial Intelligence and remote sensing for archaeological research", *Remote Sens*., 14, pp. 1-23. https://doi.org/10.3390/rs14236000.

Assael, Y., Sommerschield, Th. and Prag, J. (2019), "Restoring ancient text using deep learning: A case study on Greek epigraphy", in

Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong, China, November 3–7, 2019, Hong Kong, Association for Computational Linguistics, pp. 6368-6375. https://doi.org/10.18653/v1/D19-1668.

Assael, Y., Sommerschield, Th., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J. and de Freitas, N. (2020), "Restoring and attributing ancient texts using deep neural networks", *Nature*, 603, pp. 280-283. https://doi.org/10.1038/s41586-022-04448-z.

Balla, A., Pavlogeorgatos, G., Tsiafakis, D. and Pavlidis, G. (2012a), "Locating Macedonian tombs using predictive modelling", *JCH*, 14, pp. 403-410. http://dx.doi.org/10.1016/j.culher.2012.10.011

Balla, A., Pavlogeorgatos, G., Tsiafakis, D. and Pavlidis, G. (2012b), "Predicting Macedonian tombs' locations using GIS, predictive modeling and fuzzy logic", en *CAA 2012*, Southampton, University of Southampton - CAA International.

Balla, A., Pavlogeorgatos, G., Tsiafakis, D. and Pavlidis, G. (2013), "Modelling archaeological and geospatial information for burial site prediction, identification and management", *IJHDE*, 2, pp. 585–609. https://doi.org/10.1260/2047-4970.2.4.585

Bamman, D. and Burns, P.J. (2020), "Latin BERT: A contextual language model for Classical Philology", *arXiv*, pp. 1-14. https://doi.org/10.48550/arXiv.2009.10053

Bengio, Y. (2009), "Learning Deep Architectures for AI", *Found. Trends Mach. Learn*., 2, pp. 1-127. https://doi.org/10.1561/2200000006.

Bengio, Y., LeCun, Y. and Hinton, G. (2021), "How can neural networks learn the rich internal representations required for difficult tasks such as recognizing objects or understanding language?", *Commun. ACM*, 64, pp. 58-65. https://doi.org/10.1145/3448250.

Chankyu, L. and Hyeongjoo, K. (2020), "Fundamentos de las Humanidades de la Inteligencia Artificial", *JAHR - Revista Europea de Bioética*, 1(11), pp. 189-207.

Chapinal-Heras, D. and Díaz-Sánchez, C. (2024), "A review of AI applications in human sciences research", *DAACH*, 32, p. e00323. https://doi.org/10.1016/j.daach.2024.e00323.

Churchill, R. and Singh, L. (2022), "The evolution of Topic Modeling", *ACM Comput. Surv.*, 54(10), pp. 215:1-35. https://doi.org/10.1145/3507900.

Cowen-Breen, C., Brooks, C., Haubold, J. and Graziosi, B. (2023), "Logion: Machine Learning for Greek Philology", *arXiv*, pp. 2305.01099.

Davis, D.S., Gaspari, G., Lipo, C.P. and Sanger, M.C. (2021), "Deep learning reveals extent of Archaic Native American shell-ring building practices", *J. Archaeol. Sci.*, 132, pp. e105433. https://doi.org/10.1016/j.jas.2021.105433.

Demir, N., Boyoğlu, C.S. and Kayikci, D. (2023), "A web scrapping and AI approach for archeologists to analyze the ancient cities", *CUHES*, 4, pp. 1-8. https://doi.org/10.58598/cuhes.1213426

Deravignone, L. and Jánica, G.M. (2006), "Artificial Neural Networks in archaeology", *Archeol. Calc.,* 17, pp.121-136.

Devlin, J., Chang, M.-W, Lee, K. and Toutanova, K. (2019), "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", in *Proceedings of NAACL-HLT 2019. Minneapolis, Minnesota, June 2 - June 7, 2019*, Minneapolis, Association for Computational Linguistics, pp. 4171-4186. https://doi.org/10.18653/v1/N19-1423.

Díaz-Sánchez, C. and Chapinal-Heras, D. (2023), "Use of Open Access AI in teaching classical antiquity. A methodological proposal", *JCT*, 25(49), pp. 17-21, https://doi.org/10.1017/S2058631023000739

García-Serrano, A. and Menta Garuz, A. (2022), "La inteligencia artificial en las Humanidades Digitales: Dos experiencias con corpus digitales", *Rev. Humanid. Digit.,* 7, pp. 19-39. https://doi.org/10.5944/rhd.vol.7.2022.30928.

Griffin, S.M. (2023), "Special issue: Epigraphy and Paleography: Bringing records from the distant past to the present", *Int. Jour. Dig. Lib.,* 24, pp. 77-85. https://doi.org/10.1007/s00799-023-00371-4

Hatzinikolaou, E., Hatzichristos, Th., Siolas, A. and Mantzourani, E. (2003), "Predicting archaeological site locations using G.I.S. and fuzzy logic", in M. Doerr and A. Sarris (eds.), *The Digital Heritage of Archaeology. CAA2002. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 30th CAA Conference, Heraklion, Crete, April 2002*, Athens, Hellenic Ministry of Culture, pp. 169-176.

Iyer, A. and Franklin, M. (2022), "AI-Powered Archaeology: Determining the origin culture of various ancient artifacts using Machine Learning", *JSR*, 11, pp. 1-9. https://doi.org/10.47611/jsrhs.v11i1.2465.

Jamil, A.H., Yakub, F., Azizul Azizan, A., Roslan, S.A., Zaki, S.A. and Ahmad, S.A. (2022), "A review on Deep Learning application for detection of archaeological structures", *J. Adv. Res. Appl. Sci. Eng. Technol.,* 26, pp. 7-14. https://doi.org/10.37934/araset.26.1.714.

Jarrahi, M.H. (2018), "Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making", *Bus. Horiz*., 61, pp. 577-586. https://doi.org/10.1016/j.bushor.2018.03.007.

Keersmaekers, A. (2021), "The GLAUx corpus: methodological issues in designing a long-term, diverse, multilayered corpus of Ancient Greek", in *Proceedings of the 2nd International Workshop on Computational Approaches to Historical Language Change 2021*, Association for Computational Linguistics, pp. 39–50. https://doi.org/10.18653/v1/2021.lchange-1.6

Khan, S., Rahmani, H., Shah, S.A.A. and Bennamoun, M. (2022), *A Guide to Convolutional Neural Networks for Computer Vision*, Cham, Springer. https://doi.org/10.1007/978-3-031-01821-3.

Kolve, E., Mottaghi, R., Han, W., VanderBilt, E., Weihs, L., Herrasti, A., Deitke, M., Ehsani, K., Gordon, D., Zhu, Y., Kembhavi, A., Gupta, A. and Farhadi, A. (2022), "AI2-THOR: An interactive 3D environment for visual AI", *ArXiv*. https://doi.org/10.48550/arXiv.1712.05474.

Koentges, Th. (2020), "Measuring Philosophy in the first thousand years of Greek literature", *Digit. Class. Online*, 6, pp. 1-23. https://doi.org/10.11588/dco.2020.2.73197.

Küçükdemirci, M. and Sarris, A. (2020), "Deep learning based automated analysis of archaeo-geophysical images", *Archaeol. Prosp*., 27, pp. 108–118. https://doi.org/10.1002/arp.1763

Küçükdemirci, M. and Sarris, A. (2022), "GPR data processing and interpretation based on Artificial Intelligence approaches: Future perspectives for archaeological prospection", *Remote Sens*., 14, p. 3377. https://doi.org/10.3390/rs14143377

LeCun, Y., Bengio, Y. and Hinton, G. (2015), "Deep learning", *Nature*, 521, pp. 436-444. https://doi.org/10.1038/nature14539.

Lee, Ch. and Kim, H. (2020), "Groundwork of Artificial Intelligence Humanities", *JAHR - European Journal of Bioethics*, 1(11), pp. 189-207. https://doi.org/10.21860/j.11.1.10

Liuzzo, P.M. (2019), "EAGLE continued: IDEA. The International Digital Epigraphy Association", in A. De Santis and I. Rossi (eds.), Crossing Experiences in Digital Epigraphy. From Practice to Discipline, Warsaw, De Gruyter, pp. 216-230. https://doi.org/10.1515/9783110607208-018

Luo, J., Cao, Y. and Barzilay, R. (2019), "Neural decipherment via minimum-cost flow: from Ugaritic to Linear B", in *Proceedings of the 57th Annual Meeting of the Association for Computational*

Linguistics, Florence, Association for Computational Linguistics, pp. 3146-3155. https://doi.org/10.18653/v1/P19-1303.

Makridakis, S. (2017), "The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms", *Futures*, 90, pp. 46-60. https://doi.org/10.1016/j.futures.2017.03.006.

Mantovan, L. and Nanni, L. (2020), "The computerization of Archaeology: Survey on Artificial Intelligence techniques", *SN Computer Science*, 1, p. 267. https://doi.org/10.1007/s42979-020-00286-w

Maravelia, A., Bontozoglou, N., Kalogerakou, K., Couvaris, C. M. and Geroulanos, S. (2019), "Application of smart informatics in Egyptology: The Athens mummy project as an example of effective interdisciplinarity", *Shodoznavstvo*, 84, pp. 127-162. https://doi.org/10.15407/skhodoznavstvo2019.84.127

Soroush, M., Mehrtash, A., Khazraee, E. and Ur, J. (2020), "Deep Learning in archaeological remote sensing: Automated Qanat detection in the Kurdistan rRegion of Iraq", *Remote Sens*., 12, p. e500. https://doi.org/10.3390/rs12030500.

Orengo, H.A. and Garcia-Molsosa, A. (2019), "A brave new world for archaeological survey: Automated machine learning-based potsherd detection using high-resolution drone imagery", *J. Archaeol. Sci*., 112, p. e105013. https://doi.org/10.1016/j.jas.2019.105013.

Pagé-Perron, E., Sukhareva, M., Khait, I. and Chiarcos, Ch. (2017), "Machine translation and automated analysis of the sumerian language", in *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Vancouver, Association for Computational Linguistics, pp. 10-16. https://doi.org/10.18653/v1/W17-2202.

Panagopoulos, M., Papaodysseus, C., Rousopoulos, P., Dafi, D. and Tracy, S.V. (2009), "Automatic writer identification of ancient Greek

inscriptions", *IEEE TPAMI*, 31, 1404-1414. https://doi.org/10.1109/TPAMI.2008.201.

Papaodysseus, C., Rousopoulos, P., Arabadjis. D., Panopoulou, F. and Panagopoulos, M. (2010), "Handwriting automatic classification: Application to ancient Greek inscriptions", in *2010 International Conference on Autonomous and Intelligent Systems, AIS 2010*, Povoa de Varzim, IEEE, pp. 1-6. https://doi.org/10.1109/AIS.2010.5547045

Paraskevas, K., Mariolis, I., Giouvanis, G., Peleka, G., Zampokas, G. and Tzovaras, D. (2023), "Underwater detection of ancient pottery sherds using Deep Learning", *IJCI*, 12, pp. 1-8. https://doi.org/10.5121/ijci.2023.120601

Parker, C. S., Parsons, S., Bandy, J., Chapman, C., Coppens, F., and Seales, W.B. (2019), "From invisibility to readability: Recovering the ink of Herculaneum", *Plos One*, 14(5), p. e0215775. https://doi.org/10.1371/journal.pone.0215775

Pavlidis, G., Ben-Yosef, E. and Jones, I.W.N. (2023), "From digital recording to advanced AI applications in archaeology and cultural heritage", in E. Ben-Yosef, Erez and I.W.N. Jones (eds.), *'And in Length of Days Understanding' (Job 12:12). Essays on Archaeology in the Eastern Mediterranean and Beyond in Honor of Thomas E. Levy*, Cham, Springer, pp. 1627-1656. https://doi.org/10.1007/978-3-031-27330-8_69

Ravichandiran, S. (2021), *Getting Started with Google BERT: Build and Train State-of-the-art Natural Language Processing Models Using BERT*, Birmingham, Packt Publishing.

Rousopoulos, P., Panagopoulos, M., Papaodysseus, C., Panopoulou, F., Arabadjis. D., Tracy, Stephen, Giannopoulos, F. and Zannos, S. (2011), "A new approach for ancient inscriptions' writer identification", in *2011 17th International Conference on Digital Signal Processing (DSP)*, Corfu, IEEE, pp. 1-6. https://doi.org/10.1109/ICDSP.2011.6004966

Sharafi, S., Fouladvand, S., Simpson, I. and Álvarez, J.A.B. (2016), "Application of pattern recognition in detection of buried archaeological sites based on analysing environmental variables, Khorramabad Plain, West Iran", *J. Archaeol. Sci. Rep*., 8, pp. 206-215. https://doi.org/10.1016/j.jasrep.2016.06.024.

Smith, D.A., Rydberg-Cox, J.A. and Crane, G.R. (2000), "The Perseus Project: A digital library for the Humanities", *LLC*, 15, 15-25.

Sommerschield, Th., Assael, Y., Pavlopoulos, J., Stefanak, V., Senior, A., Dyer, Ch., Bodel, J., Prag, J., Androutsopoulos, I. and de Freitas, N. (2023), "Machine Learning for ancient languages: A survey", *Comput. Linguist*., 49, 1-44. https://doi.org/10.1162/coli_a_00481.

Tracy, S.V., Papaodysseus, C., Rousopoulos, P., Panagopoulos, M., Fragoulis, D., Dafi, D. and Panagopoulos, Th. (2007), "Identifying hands on ancient Athenian inscriptions: First steps towards a digital approach", *Archaeometry*, 49, pp. 749-764. https://doi.org/10.1111/j.1475-4754.2007.00333.x.

UNESCO (2019), *Beijing Consensus on Artificial Intelligence and Education*, Beijin, UNESCOg. https://unesdoc.unesco.org/ark:/48223/pf0000368303.

Vaishya, R., Javaid, M., Khan, I.H., Haleem, A. (2020), "Artificial Intelligence (AI) applications for COVID-19 pandemic", *Diabetes Metab. Syndr.: Clin. Res*., pp. 14, 337-339. https://doi.org/10.1016/j.dsx.2020.04.012.

Wayenberg, J., Capriotti, M. (2014), "Investigating Greek painted iconology by using Artificial neural networks. Maenads and satyrs on Athenian red-figure pottery", in M. Ramazzotti (ed.), *ARCHEOSEMA. Artificial Adaptive for the Analysis of Complex Phenomena. Collected Papers in Honour of David Leonard Clarke, Archeologia e Calcolatori*, Firenze, All'Insegna del Giglio, pp. 263-290.