



A review of AI applications in Human Sciences research

Diego Chapinal-Heras^{a,*}, Carlos Díaz-Sánchez^b

^a Universidad Autónoma de Madrid (Autonomous University of Madrid), C. Francisco Tomás y Valiente, 1, 28049, Madrid, Spain

^b Universidad Complutense de Madrid (Complutense University of Madrid), C. Profesor Aranguren s/n, 28040, Madrid, Spain

ARTICLE INFO

Keywords:

AI technology
Humanities
Human sciences
Digital Humanities
History
Palaeography
Epigraphy
Archaeology
History of Art

ABSTRACT

This paper compiles the most relevant contributions of AI technology in the field of Human Sciences. It reveals the development of different applications and the software and techniques used. The main purpose is to show the major breakthroughs in Humanities research and, at the same time, to highlight the scarcity of study cases. The implementation of AI offers many possibilities that are gradually being incorporated. This general state-of-the-art review aims to illustrate how our area of knowledge can benefit from the different approaches that have been undertaken, as well as their potential for future developments.

1. Introduction

The aim of this paper is to review the literature on Artificial Intelligence (AI) in the sphere of the Human Sciences, primarily History, Literature and other connected fields of research, such as Palaeography, Epigraphy and Archaeology. AI has been applied to a larger extent in other knowledge areas for varying purposes. A few examples are the improvement of business operations (Makridakis, 2017; Jarrahi, 2018), the development of virtual sceneries (Kolve et al., 2022) and healthcare (Vaishya et al., 2020). UNESCO included the integration of AI in education as part of the 2030 Agenda for Sustainable Development (UNESCO, 2019). In the Humanities, however, the process is slower. In this paper we will see the most relevant cases in which AI has been implemented or is gradually being applied.

The meaning, functions and uses of AI are diverse. Fundamentally, its purpose is to enable machines to develop the ability to interpret external data and perform that task accurately, to automatically learn from that practice, and ultimately to use that knowledge in specific assignments (García-Serrano and Menta Garuz, 2022). The way to achieve this is through the recompilation, classification and management of large volumes of data. By using a series of algorithms, computers are not only able to correctly undertake tasks, but also to interact with agents that use this resource or tool, as well as to learn how to improve their efficiency (Bengio 2009).

The development of AI has gone through different phases, progressively incorporating more resources and increasing in potential. As will be shown, this process has had a direct impact on its application in the Human Sciences. For example, from text processing for the identification of terms and being able to transliterate, we have reached the possibility of partially or totally restoring fragmentary inscriptions. The software trains and learns. Similarly, beginning with reducing the amount of work, AI is today able to approach the skills of professionals in different fields. This point has been reached mainly thanks to the crucial progress undergone in Deep Learning, the branch of AI that attempts to allow machines to think and learn in the same way as the human brain. To do this, scientists make use of the model of neural networks that can simultaneously process data at different levels (Bengio et al., 2021). Within this field, Convolutional Neural Networks (CNN) are especially useful for virtual analysis. CNN models can be used for both supervised and unsupervised learning paradigms. In the first mechanism, the input to the system and the desired outputs (true labels) are known and the model learns a mapping between the two. In the unsupervised learning mechanism, on the other hand, the true labels for a given set of inputs are not known and the model aims to estimate the underlying distribution of the input data samples (Khan et al., 2022).

Another remarkable parallel improvement has been Natural Language Processing (NLP), which is able to perform tasks such as text restoration. NLP consists of the development of computational

* Corresponding author. Departamento Historia Antigua, Medieval, Paleografía y Diplomática, Facultad de Filosofía y Letras, Universidad Autónoma de Madrid, C. Francisco Tomás y Valiente, 1, 28049, Madrid, Spain.

E-mail address: diego.chapinal@uam.es (D. Chapinal-Heras).

<https://doi.org/10.1016/j.daach.2023.e00288>

Received 30 March 2023; Received in revised form 8 July 2023; Accepted 21 August 2023

Available online 22 August 2023

2212-0548/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

mechanisms that boost communication between people and machines by means of language. This includes the use of algorithms and different methods through which the machine automatically learns (Pagé-Perron et al., 2017).

It is worth pointing out that the application of AI in the Humanities does not purport to replace the human task. Actually, all the studies consulted in this paper coincide in that the main purpose of AI is to act as a resource to allow research to be undertaken more quickly and efficiently. After all, computers are able to process large amounts of information much more quickly than a human professional in any specific topic. Therefore, AI must be seen as a useful tool in Human Sciences research, instead of an obstacle or even a threat to our work. In fact, the most frequent procedure is for a specialist to review the task undertaken by AI to check it for mistakes, given that it is very uncommon for AI to obtain 100% accuracy. The reviewer then corrects whatever is necessary. This review process implies a less time-consuming activity, as it avoids the need to carry out the entire task. In fields where it is necessary to process a large volume of data, the usefulness and benefits of AI are undeniable.

The use of these new methodologies and resources is essential for all knowledge areas. If Makridakis (2017) is right, in the course of this generation, the progress with AI will be such that we will be speaking of a technological revolution greater even than the Industrial Revolution. With this in mind, it is crucial to promote AI in the Humanities. In recent decades, the development of digital technology has entailed a transformation in a large part of the different branches of our field. It is therefore a process that is not going to end any time soon. The literature mentioned below shows the different ways we can apply AI in several areas of research.

2. Methodology

The search for publications regarding the topics examined in this paper focused on two methods. Firstly, the use of bibliographic repositories, mainly Semantic Scholar and Google Scholar. In these cases, we introduced specific commands: “AI Palaeography”, “AI Epigraphy”, “AI Archaeology”, etc. The second technique was to compile more literature from the most recent publications previously found. In this way, it was possible to gather and select the most important contributions in the Human Sciences.

We classified the literature according to wide fields of study, beginning with those to which remarkable contributions have been made and in which the stage of development regarding the application of AI is more advanced. At the same time, within each topic group we attempted to follow a chronological order, thus showing the gradual incorporation of new methods and possibilities. However, in certain cases two or more closely connected studies are presented together, as they reflect the continuation of the research into a specific aspect. In this way, we can see different cases involving Literature, Palaeography, Epigraphy, Archaeology and the History of Art.

3. AI in the human sciences

AI in general and NLP and data analysis in particular have been the most frequent IT approaches in the last two decades in Digital Humanities projects (García-Serrano y Menta Garuz, 2022). The combination of different methods and techniques has allowed us to obtain valuable results. Without these technological developments and using just traditional methodology, the results would have taken several decades to obtain. To a large extent this is due to the fact that the amount of data these study cases have worked with is such that it would have required the contribution of many more individuals with expertise in those fields.

In Literature, experimentation with AI has entailed the development of new application areas. In 2013, a team consisting of DiMaggio et al. (2013) used different algorithms to enable AI to analyse discourses to identify patterns or interpretative frameworks in the African press

between 1986 and 1991. This reduced the time required to analyse a number of works to recognise the ideological trends during that period. Another noteworthy study is that of Schöch (2017) focusing on French literary production during that country's Classical Age and Enlightenment. This research used an unsupervised method within the working sphere of NLP known as Topic Modelling. This approach allows theme patterns and trends to be found in entire texts. In order to undertake this study, the author compiled 890 theatrical plays from between 1610 and 1810. The semantic analysis on different levels yielded acceptable results with a relatively high degree of accuracy and the research revealed different topics. As the author explains, however, the classification of these topics is still interpretable. It was also possible to verify that each subgenre had certain distinct topics.

More recently, Köntges (2020) attempted to identify conceptual elements in philosophical schools of thought that are repeated, reused or mentioned in the first thousand years of ancient Greek literature. The technique used in this study was Latent Dirichlet Allocation (LDA) topic modelling, which allowed document vectors to be produced based on the structural and thematic content of each analysed document. The open-access program used was (*Meletē*) *Tōpan* created by Köntges. In this case, the examined documentation contained almost 30 million words. The added difficulty with respect to the previous publication mentioned is that in this case the language processed was ancient Greek, which is more problematic in terms of its analysis. The results were very positive and showed that AI is able to extract references that contain aspects of the philosophical thought of that period.

The analysis and processing of handwritten documents is another fruitful way of using AI in literary and historical research. To our knowledge, the first palaeographic research of this kind was a project undertaken at the University of Pisa beginning in 1999. A team of postgraduate students led by Professors A. Sperduti and A. Starita created the System for Palaeographic Inspections (SPI) software, aimed at recognising similarities among the letters contained in the manuscripts. This technique allows different styles of handwriting to be identified, as well the evolution of the writing, which in turn helps to date documents to a higher degree of accuracy. Unfortunately, the project remained unfinished (Ciula, 2005). However, this approach laid the foundations for what is known as Digital Palaeography or what Wolf et al. (2011) called “computerised palaeography”. They defined this concept as “digital tools that furnish the analysis of a human palaeographer with large-scale capabilities and assist with evidence-based inference”. Within this trend, it is very useful to apply a stylometric analysis, the IT technique based on the comparison of texts according to their scriptural uses (Cuéllar, 2023).

From this point on, there are many examples of practical applications of AI in examining the content of manuscripts that go beyond mere text compilations. We can see that in the study by Batjargal et al. (2010) focusing on almost 50,000 manuscripts from the Traditional Mongolian Script Digital Library. This study allowed the documentation to be rendered and the data to be searched for in both modern Mongolian (Cyrillic) and traditional Mongolian scripts with a high degree of success. Also in Asia, Azmi et al. (2011) applied the SPI software developed at the University of Pisa to work with 7789 Malaysian manuscripts. The purpose of this study was to identify the dating and geographical location of each document based on an analysis of the writing style. Their work included the decoration of the pages with special materials –known as illumination– given that such elements contain spatial and temporal information.

Regarding Spanish, the first study is that of Meza-Lovn (2012), who presented the Handwritten Text Recognition (HTR) system. This method transcribes historical documents through graphic representation and the use of Support Vector Machine techniques to train the program. In the case of modern Italian literature, Tuzzi and Cortelazzo (2018) invested their efforts in the examination of 150 works from 40 different authors in an attempt to work out who was behind the pseudonym Elena Ferrante. To do this, they used NLP and undertook data and stylometric

analyses. They reached the conclusion that the famous novelist Starnone seems to have had some connection, although it is not possible to confirm whether that author was the very same person or just had some kind of influence.

More recent advances have made remarkable qualitative progress. For instance, Kang et al. (2021) worked with the Annals of the Joseon Dynasty and the Diaries of the Royal Secretariat from Korea, which cover the period from the 14th to the early 20th centuries. This study involves dealing with 50 and 243 million characters, respectively. The multi-task approach this team is taking aims to restore the damaged parts of the documents and to translate them into modern language. The keystone of this research is that AI learning develops through a self-attention mechanism, which is augmented as it processes more documents.

Another use of AI in palaeography is to identify the authorship of manuscripts that are preserved as “anonymous”. This approach involves Optical Character Recognition (OCR) and stylometric analysis. Very recently, a Spanish team formed by Cuéllar and Vega García-Luengos (2023) was able to link a theatrical comedy, *La Francesa Laura* (French Laura), to Lope de Vega. The manuscript is preserved in the Spanish National Library. With the use of *Transkribus*, a text recognition tool, and models developed to identify texts, the researchers were able to transcribe and automatically modernise theatre printings and manuscripts from the Spanish Golden Age of Literature. The program was trained with more than three million words from the literary genre of that period.

Some progress is also worth mentioning in Classical Antiquity. Since the establishment of the Perseus Project at the end of the 20th century, numerous studies have taken advantage of the possibilities offered by IT to work with texts and the different elements related to them. AI use has been incorporated gradually within the framework of the implementation of new technologies. The development in 2018 of BERT (Bidirectional Encoder Representation from Transformer) by Google opened up new possibilities. The main function of this technique is the pre-training of NLP to allow the system to improve its analysis and interpretation of user searches. The key point of BERT, which led to a major breakthrough in NLP, is that it is a context-based embedding model. This means that each word is examined considering the rest of terms that the sentence contains, thus providing a more accurate understanding, especially when dealing with words that may have more than one meaning, depending on the context (Ravichandiran, 2021).

The BERT technique has been used by many researchers to create specific applications in certain languages. A good example is Latin-BERT built by Bamman and Burns (2020) that uses Latin sources ranging from 200 BCE to the present day. Although Classical Latin is no longer spoken, its use and evolution in different contexts over such a long period have made it possible to train BERT with 642.7 million tokens. Thanks to this, the software can emend fragmentary texts. More specifically, the model provides estimates of word probabilities, allowing the researcher to check the most likely and accurate one in each case.

The initial steps in the field of Classical Epigraphy took place at the beginning of the 21st century with a very specific study case: the Vindolanda texts. That archaeological site, one of the *castra* situated along Hadrian's Wall in Roman Britain, has yielded a large number of ink tablets. These documents contain both military and personal information about the people who lived there. M. Terras inaugurated the application of computer techniques to analyse the texts in her PhD presented in 2002 and published as a book four years later (Terras, 2006). This first approach consisted of processing the documents, enabling computers to read them in a similar way to papyrologists, thus assisting them. Further work was undertaken later by Roued-Cunliffe (2010), who used the Decision Support System (DSS) web application DUGA and developed a knowledge-based web service called APOLLO. The building of this service also required the reformatting of the Vindolanda dataset into EpiDoc standard XML. This expanded the possibilities of transcribing and reading the tablets, including restoring lost

letters or words with low and high level perceptss respectively.

In parallel, epigraphy has seen important progress in the last two decades in terms of analysis with IT resources. One of the first studies was by Tracy et al. (2007), who focused their efforts on the identification of the hands of letter cutters in a compilation of inscriptions from ancient Athens. This first study was conducted with just six documents, but it was enough to demonstrate the potential of the technology, which converts all the information contained in the letter strokes into mathematical formulae. Technically, this approach does not belong to the AI sphere, but to the development of mathematical operations to process such sources. It does not include a specific software to enable learning through training, although we can consider it a relevant step towards the application of AI. This methodology was quickly expanded and improved by a team consisting of some of the authors of the previous contribution. In this case, the study was of 24 Athenian inscriptions, and it was possible to recognise six different writers (Panagopoulos et al., 2009).

From this point on, AI has been used to research epigraphical materials from different parts of the world. Soumya and Hemantha Kumar (2014) used a Random Forest Classification to examine documents written in Kannada, one of the ancient languages of India, in an attempt to establish their chronology. Can et al. (2016) studied Mayan glyphs, focusing on their visual analysis. More specifically, they worked on the effects of model parameters of single-hidden-layer Sparse Autoencoder (SA) to test whether they were capable of comparable results to those of traditional descriptors. For the Sumerian civilisation, Pagé-Perron et al. (2017) undertook a project based on a combination of NLP unsupervised methods and Neural Machine Translation (NMT). The aim was to help translate the enormous corpus of Sumerian documents, as we currently know the contents of just 10% of all the preserved inscriptions. Luo et al. (2019) did something similar with their application of a neural decipherment algorithm to Near Eastern Ugarit texts and Mycenaean Greek Linear B.

Greek and Latin inscriptions were the subject of the research carried out by Amato et al. (2016). That project was developed within the framework of the EAGLE network, which compiled most of the epigraphs from those cultures. The team focused on Fisher Vectors (FVs) for encoding local descriptors and various CNN representations, including a combination of both FVs and CNNs. With these approaches, they performed a visual similarity search among all the images in the dataset to enable the system to recognise the object in a query image.

The most recent contributions in this field belong to the group led by Assael. They first developed PYTHIA, a fully automated deep learning model trained to restore the text of ancient Greek inscriptions by predicting the character sequences comprising the hypothesised restorations. That was supplemented by PHI-ML, a dataset of machine actionable text consisting of more than 3.2 million words (Assael et al., 2019). Recently, this research has been expanded and reinforced with the development of Ithaca, software that applies Deep Neural Networks (DNNs) to improving the analysis and restoration of epigraphical documents written in ancient Greek. Like PYTHIA and other prior studies in this sphere, Ithaca began on the basis of the existence of formulae, patterns and expressions that tend to repeat. By means of textual and contextual parallelism, the system is able to identify with a high degree of accuracy the geographic and temporal location of these inscriptions. This entails the generation of a dataset of machine-actionable epigraphic text containing 76,608 texts. Once the results have been obtained, the researchers' task is to review them and check that the reconstructions suggested by the program are solid or whether they need to be modified (Assael et al., 2020).

Archaeology has also seen the development of some practical applications of AI. They can be summarised in deep learning and data processing and analysis, mainly for telematic prospections and remote sensing (Sharafi et al., 2016; Soroush et al., 2020; Bundzel et al., 2020; Davis et al., 2021). The use of drones is very common in this area (Argyrou and Agapiou 2022). There is software that uses DNNs for the

identification of artifacts and structures during telematic surveys (Trier et al., 2019). These systems implement self-learning and deep learning algorithms to process and classify artifacts discovered by remote sensing, making it easier to automatically identify and categorise them (Jamil et al., 2022).

Orengo and Garcia-Molsosa (2019) introduced AI self-learning algorithms into unmanned aircraft that use high definition remote sensing systems to survey areas more quickly and are able to recognise features on the surface or just below it. This task allows savings to be made on human and financial resources. Similarly, Caspari and Crespo (2019) used the same technology to remotely detect necropolises and to facilitate the analytical processes and the search for traces of burials. Along the same lines, Berganzo-Besga et al. (2021) made use of self-learning algorithms in spectral remote sensing and LiDAR for the automatic detection of funerary tumuli in northwestern Iberia. Another relevant application is that of the software developed by Domínguez-Rodrigo et al. (2020), who used Big Data to study different bone marks. The AI was programmed to identify each sort of mark and help to interpret them.

In the field of History of Art, technical progress based on image scanning and virtualisation has allowed the generation of Big Visual Data that AI can work with. Using algorithms, it has been possible to recognise works of doubtful or anonymous authorship (Gefen et al., 2021). Other remarkable research, which combines historical and art-historical study, is that of DIMH (*El Dibujante Ingeniero al servicio de la Monarquía Hispánica. Siglos XVI-XVIII/The Engineer Draughtsman of the Hispanic Monarchy. Sixteenth to Eighteenth Centuries*). Its aim is to develop tools to undertake the tasks of semantic annotation, searching for information, extracting hidden relationships in texts, and the visualisation of the results (García-Serrano and Castellanos González, 2016).

4. Discussion and conclusions

Despite the advantages and benefits AI offers data processing and analysis, its application in the Human Sciences is relatively rare. There are several factors that might explain this, such as the financial cost, the complexity of the processes and algorithms related to AI programming, and the databases used. Changing traditional research dynamics in the Humanities is a slow process. In other areas of knowledge, the application of new technologies represents the very core of most of the progress. In the case of areas such as historical and literary studies, however, the situation is more complicated. Nevertheless, new tools are gradually being incorporated. With them, we can process and examine the available data more quickly and accurately.

This paper has demonstrated the potential of AI in Human Sciences research. It has described several study cases in which experimentation with the different methods and techniques developed in recent years has contributed to a major breakthrough in data analysis. A task that would have required an expert in the field several years to undertake can be carried out by trained AI software in just a few weeks or even days. The role of AI as a tool designed to aid the researcher –and not to replace them– strengthens the usefulness of this type of work. AI is not a technology that is able to examine a body of data automatically with no human intervention in any of the steps in the process. The figure of the academic is essential for reviewing the results produced by the software in order to modify, complete or expand the results.

There are areas in the Humanities that to date have received relatively little attention regarding the use of AI. This is the case of History of Art and Archaeology studies. In the latter, most of the cited contributions focus on the remote detection of material remains during surveys. The combination of this technique and GIS improves the efficacy of the identification of structures and architectural and urban development patterns. This means an excavation can be initiated with a much better knowledge of the area. Nevertheless, there are many potential applications of AI in archaeological studies that still remain in a mere embryonic state. This is the case, for example, of the quantitative and

qualitative analysis of finds already excavated and inventoried. Undoubtedly, there are many possibilities here, such as seeking common or divergent patterns regarding the procurement of raw materials to manufacture certain objects, the different functions these artifacts may have had, or their evolution over time. This kind of automated study would complement the research undertaken until now by experts, allowing a larger number of finds to be covered in less time.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Diego Chapinal-Heras reports writing assistance was provided by Spanish Ministry of Science and Innovation of Spain by means of the Postdoctoral Fellowship Ramón y Cajal. Diego Chapinal-Heras reports a relationship with Spanish Ministry of Science and Innovation that includes: postdoctoral fellowship with funding grants.

References

- Amato, G., Falchi, F., Vadicamo, L., 2016. Visual recognition of ancient inscriptions using convolutional neural network and Fisher vector. *Journal on Computing and Cultural Heritage* 9 (21), 1–24. <https://dl.acm.org/doi/10.1145/2964911>.
- Argyrou, A., Agapiou, A., 2022. A Review of Artificial Intelligence and remote sensing for archaeological research. *Rem. Sens.* 14, 1–23. <https://doi.org/10.3390/rs14236000>.
- Assael, Y., Sommerschild, T., 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. Association for Computational Linguistics, Hong Kong, pp. 6368–6375.
- Assael, Y., Sommerschild, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androustopoulos, I., Prag, J., de Freitas, N., 2020. Restoring and attributing ancient texts using deep neural networks. *Nature* 603, 280–283. <https://doi.org/10.1038/s41586-022-04448-z>.
- Azmi, M.S., Omar, K., Nasrudin, M.F., Muda, A.K., Abdullah, A., 2011. Digital paleography: using the digital representation of Jawi manuscripts to support paleographic analysis. In: *2011 International Conference on Pattern Analysis and Intelligence Robotics*. IEEE, Putrajaya, pp. 71–77.
- Bamman, D., Burns, P.J., 2020. Latin BERT: a contextual language model for Classical Philology. *arXiv* 1–14. <https://doi.org/10.48550/arXiv.2009.10053>.
- Batjargal, B., Khaltarkhuu, G., Fuminori, K., Fuminori, M.A., 2010. Ancient-to-modern information retrieval for digital collections of traditional Mongolian script. In: *International Conference on Asian Digital Libraries*. ICADL, Gold Coast, pp. 25–28.
- Bengio, Y., 2009. Learning deep architectures for AI. *Foundations and Trends in Machine Learning* 2, 1–127. <https://doi.org/10.1561/2200000006>.
- Bengio, Y., Lecun, Y., 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, 58–65.
- Berganzo-Besga, I., Orengo, H.A., Lumberras, F., Carrero-Pazos, M., Fonte, J., Vilas-Estévez, B., 2021. Hybrid MSRM-based DeepLearning and MultitemporalSentinel 2-based machine learning algorithm detects near 10k archaeological tumuli in northwestern iberia. *Rem. Sens.* 13, 4181. <https://doi.org/10.3390/rs13204181>.
- Bundzel, M., Jascur, M., Kováč, M., Lieskovský, T., Sincák, P., Tkáčik, T., 2020. Semantic segmentation of airborne LiDAR data in maya Archaeology. *Rem. Sens.* 12, 3685. <https://doi.org/10.3390/rs12223685>.
- Can, G., Odobez, J.-M., Gatica-Perez, D., 2016. Evaluating shape representations for maya glyph classification. *Journal on Computing and Cultural Heritage* 9. <https://doi.org/10.1145/2905369>.
- Caspari, G., Crespo, P., 2019. Convolutional neural networks for archaeological site detection-finding “princely” tombs. *J. Archaeol. Sci.* 110, 104998. <https://doi.org/10.1016/j.jas.2019.104998>.
- Ciula, A., 2005. Digital palaeography: using the digital representation of medieval script to support palaeographic analysis. *Digit. Mediev.* 1.
- Cuellar, A., 2023. Cronología y estilometría: Datación automática de comedias de Lope de Vega, vol. XXIX. *Anuario Lope de Vega*, pp. 97–130.
- Cuellar, A., Vega García-Luengos, G., 2023. La Francesa Laura. El hallazgo de una nueva comedia del Lope de Vega último, vol. XXIX. *Anuario Lope de Vega*, pp. 131–198.
- Davis, D.S., Gaspari, G., Lipo, C.P., Sanger, M.C., 2021. Deep learning reveals extent of Archaic Native American shell-ring building practices. *J. Archaeol. Sci.* 132, 105433. <https://doi.org/10.1016/j.jas.2021.105433>.
- DiMaggio, P., Nag, M., Blei, D., 2013. Exploiting affinities between topic modeling and thesocological perspective on culture: application to newspaper coverage of U.S. governmentarts fundig. *Poetics* 41 (5), 570–606.
- Domínguez-Rodrigo, M., Cifuentes-Alcobendas, G., Jiménez-García, B., Abellán, N., Pizarro-Monzo, M., Organista, E., Baquedano, E., 2020. Artificial intelligence provides greater accuracy in the classification of modern and ancient bone surface modifications. *Nature* 10, 18862. <https://doi.org/10.1038/s41598-021-82916-8>.
- García-Serrano, A., Castellanos González, A., 2016. Conceptualización, acceso y visibilidad de la información en el proyecto DIMH. In: *Cámara Muñoz, A. (Ed.), El*

- Dibujante Ingeniero al Servicio de La Monarquía Hispánica: Siglos XVI XVIII. Fundación Juanelo Turriano, Madrid, pp. 379–400.
- García-Serrano, A., Menta Garuz, A., 2022. La inteligencia artificial en las Humanidades Digitales: dos experiencias con corpus digitales. *Revista de Humanidades Digitales* 7, 19–39. <https://doi.org/10.5944/rhd.vol.7.2022.30928>.
- Gefen, A., Saint-Raymond, L., Venturini, T., 2021. AI for digital Humanities and computational social Sciences. In: Braunschweig, Bertrand, Ghallab, Malik (Eds.), *Reflections on Artificial Intelligence for Humanity*. Springer, Cham, pp. 191–202.
- Jamil, A.H., Yakub, F., Azizul Azizan, A., Roslan, S.A., Zaki, S.A., Ahmad, S.A., 2022. A review on Deep Learning application for detection of archaeological structures. *Journal of Advanced Research in Applied Sciences and Engineering Technology* 26, 7–14. <https://doi.org/10.37934/arasent.26.1.714>.
- Jarrahi, M.H., 2018. Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making. *Bus. Horiz.* 61, 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>.
- Kang, K., Jin, K., Yang, S., Jang, S., Choo, J., Kim, Y., 2021. Restoring and mining the records of the Joseon Dynasty via neural language modeling and machine translation. *arXiv*. <https://doi.org/10.48550/arXiv.2104.05964>.
- Khan, S., Rahmani, H., Shah, S., Afaq, A., Bennamoun, M., 2022. A Guide to Convolutional Neural Networks for Computer Vision, Synthesis Lectures on Computer Vision. Springer, Cham. Number 15.
- Kolve, E., Mottaghi, R., Han, W., VanderBilt, E., Weihs, L., Herrasti, A., Deitke, M., Ehsani, K., Gordon, D., Zhu, Y., Kembhavi, A., Gupta, A., Farhadi, A., 2022. AI2-THOR: an interactive 3D environment for visual AI. *ArXiv*. <https://doi.org/10.48550/arXiv.1712.05474>.
- Köntges, T., 2020. Measuring Philosophy in the first thousand years of Greek literature. *Digital Classics Online* 6, 1–23. <https://doi.org/10.11588/dco.2020.2.73197>.
- Luo, J., Cao, Y., 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*. Association for Computational Linguistics, Florence, pp. 3146–3155.
- Makridakis, S., 2017. The forthcoming Artificial Intelligence (AI) revolution: its impact on society and firms. *Futures* 90, 46–60. <https://doi.org/10.1016/j.futures.2017.03.006>.
- Meza-Lovn, G.L., 2012. A Graph-Based approach for transcribing ancient documents. In: *Ibero-American Conference on Artificial Intelligence*. IBERAMIA, Cartagena de Indias, pp. 210–220.
- Orengo, H.A., 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, 105013 <https://doi.org/10.1016/j.jas.2019.105013>.
- Pagé-Perron, E., Sukhareva, M., Khait, I., Chiarcos, C., 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*. Association for Computational Linguistics, Vancouver, pp. 10–16.
- Panagopoulos, M., Papaodysseus, C., Roussopoulos, P., Dafi, D., Tracy, S.V., 2009. Automatic writer identification of ancient Greek inscriptions. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 1404–1414.
- Ravichandiran, S., 2021. *Getting Started with Google BERT: Build and Train State-Of-The-Art Natural Language Processing Models Using BERT*. Packt Publishing, Birmingham.
- Roued-Cunliffe, H., 2010. Towards a decision support system for reading ancient documents. *Lit. Ling. Comput.* 25, 365–379.
- Schöch, C., 2017. Topic modeling genre: an exploration of French classical and enlightenment drama. *DHQ: Digital Humanities Quarterly* 11.
- Sharafi, S., Fouladvand, S., Simpson, I., Alvarez, 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.: Report* 8, 206–215. <https://doi.org/10.1016/j.jasrep.2016.06.024>.
- Sorosh, M., Mehrtash, A., Khazraee, E., Ur, J., 2020. Deep learning in archaeological remote sensing: automated qanat detection in the kurdistan region of Iraq. *Rem. Sens.* 12, 500. <https://doi.org/10.3390/rs12030500>.
- Soumya, A., Hemantha Kumar, G., 2014. Classification of ancient epigraphs into different periods using random forests. In: *Fifth International Conference on Signal and Image Processing*. IEEE, pp. 171–178.
- Terras, M., 2006. *Image to Interpretation: an Intelligent System to Aid Historians in Reading the Vindolanda Texts*. Oxford University Press, Oxford.
- Tracy, S.V., Papaodysseus, C., Roussopoulos, P., Panagopoulos, M., Fragoulis, D., Dafi, D., Panagopoulos, Th., 2007. Identifying hands on ancient Athenian inscriptions: first steps towards a digital approach. *Archaeometry* 49, 749–764. <https://doi.org/10.1111/j.1475-4754.2007.00333.x>.
- Trier, Ø.D., Cowley, D.C., Waldeland, A.U., 2019. Using deep neural networks on airborne laser scanning data: results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland. *Archaeol. Prospect.* 26, 165–175. <https://doi.org/10.1002/arp.1731>.
- Tuzzi, A., Cortelazzo, M., 2018. What is Elena Ferrante? A comparative analysis of a secretive bestselling Italian writer. *Digital Scholarship in the Humanities* 33, 685–702. <https://doi.org/10.1093/dlc/fqx066>.
- UNESCO, 2019. *Beijing Consensus on Artificial Intelligence and Education*. UNESCO, Beijing.
- Vaishya, R., Javaidb, M., Khanc, H., Halemb, A., 2020. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metabol. Syndr.: Clin. Res. Rev.* 14, 337–339. <https://doi.org/10.1016/j.dsx.2020.04.012>.
- Wolf, L., Potikha, L., Dershowitz, N., Shweka, R., Choueka, Y., 2011. Computerized paleography: tools for historical manuscripts. In: *18th IEEE International Conference on Image Processing*. IEEE, Brussels, pp. 3545–3548.