The Application of Artificial Intelligence Technologies in Digital Humanities: Applying to Dunhuang Culture Inheritance, Development, and Innovation

Junhan Ye
Glasgow College, University of Electronic Science and Technology of China, China
Corresponding Author: Junhan Ye, E-mail: 2019190602006@std.uestc.edu.cn

ABSTRACT
This article will focus on the relationship between Digital Humanities and Artificial Intelligence and will discuss the methodology of applying Artificial Intelligence in Digital Humanities; specifically, the article will target the applications in Dunhuang culture studying. The development of digital humanities facilitated by computers and the future research direction will be studied. The article aims to propose specific methods of applying Artificial Intelligence to Digital Humanities to facilitate the inheritance, development, and innovation of traditional culture. By sorting out the development process of digital humanities and combining it with the newest development direction of Artificial Intelligence forecasted by specialists in the area, the study will discuss the possible applications of the technologies in intangible cultural heritage studies. The methods include utilizing Smart Data to obtain structured and visual cultural heritage, implying Cross-media Intelligence to help reconstruct and propagate culture, and using Human-machine Association for the renovation and recreation of cultural heritage. Additionally, the specific methods of applying these methodologies to Dunhuang culture studying will be discussed, whose object is to realize natural language analysis, image style transfer using Smart Data, build digital museums using VR and AR, and restore murals in Mogao caves using Human-machine Association. The effects of these techniques on cultural heritage and its future development methodology will also be discussed.

KEYWORDS
Digital Humanities, Artificial Intelligence, Dunhuang Culture

ARTICLE INFORMATION
ACCEPTED: 27 August 2022 PUBLISHED: 30 August 2022 DOI: 10.32996/jcsts.2022.4.2.5

1. Introduction
Digital Humanities is a relatively new area of research. At the end of the 19th century, American physicist Thomas Corwin Mendenhall first analyzed the writing habits of several writers by measuring the words in all their works (Dai et al., 2016). Several other literary critics and historians have also used similar statistical approaches in their research (Dai et al., 2016; Zhong & Zhang, 2016). This tendency of utilizing quantitative analysis in humanities, the prehistory of Digital Humanities, raised a new subject. From paper to computer, the efficiency of data processing had greatly increased, which was the point where “humanities computing” turned to “digital humanities”. In 1949, Roberto Busa made a groundbreaking change to humanities computing, where he used a computer to make an electronic index including 11 million medieval Latin vocabulary (Hockey, 2004). After that, more researchers started to realize the advantage of computer indexing and created several vocabulary indexes for different corpora; for example, Burton (1982) made a summary of the attempts of some scholars on text indexing. More development has been made to improve the searching of certain information in the context, such as the COCOA indexing program and those discussions about coding and indexing in journals such as (Bessinger & Parrish, 1964) and forums such as ICCH and ACH. At that time, some organizations were founded to promote the utilization of computers in humanities, such as the Centre for Literary and Linguistic Computing at Cambridge University. As this area had been paid more and more attention, standardized programs and principles were developed to satisfy the need of scholars and even domestic computer users from worldwide to work together (Hockey, 2004).

Copyright: © 2022 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.
However, the essence of human and computer processing mentioned above were the same, where the key point of the research was quantizing text data, building databases of mass text, and indexing. As was said by some scholars that text is the most basic and important research object of humanities computing; in the development process of humanities computing, applications related to text resources have occupied a core position, which can be seen from a large number of related publications (Nyhan & Flinn, 2016). Moreover, a new level was reached in the electronic area of humanities, where people tried to share and obtain multimedia information in the form of pictures, audio, and video (Hockey, 2004). Until now, with the mass applications of the Internet and multimedia, producing and spreading the speed of knowledge is explosively increasing, and the interaction between different cultures and different subjects is increasing as well; therefore, a new methodology should be implied to break the limitations of independent text analysis. The new pattern, proposed in The State of the Digital Humanities: A Report and a Critique, was called “Distant Reading”, where literature is put in context to interpret and analyze, explaining the huge system of texts in terms of category factors and mode elements (Liu, 2011). The essential difference between “humanities computing” and “digital humanities” is the change from “hypothesis testing and material confirmation” by humans to “hypothesizing, observing, discovering, analyzing, testing, and hypothesizing again” by both humans and machines (Dai et al., 2016). There are 2 tendencies in Digital Humanities: the first is to turn from small, hand-entered databases to Big Data projects; the second is to lower the technical threshold for using computers to analyze texts, which will promote the combination of Digital Humanities and Artificial Intelligence because people that are not proficient in the computer can use computer tools in humanities researching (Dai et al., 2016). In our research, Digitalized Collection will be the initial step to connecting computers with the real world; after it, Data Arrangement and Analysis work as the basis for digitalization; then advanced technologies such as Visualization, Machine Learning, and Smart-data Driving Learning are needed for specific application in humanities. As a representative example of intangible cultural heritage, Dunhuang culture is a fusion of different civilizations, ages, and human subjects. With rich literature, images, audio, and arts resources, Dunhuang culture is appropriate for Digital Humanities applications. There have been many types of research by Chinese and foreign scholars studying the digitalization of Dunhuang Culture. For example, the problems of image collecting and processing with higher accuracy were discussed in (Aoki et al., 2004; Department of Aerial Survey & Institute of Archaeology, 1996; Gao, 2008; Liu, 2000; Wei et al., 1999); also, the achievements of Dunhuang literature digitalization were shown in refs (Chen & Wu, 1997; Han, 2009; Haritori & Masato, 2005; Kosei, 2002; Liu & Ye, 2017; Ma, 2006; Whitfield, 2001). However the utilization of Artificial Intelligence techniques in this area is deficient. The advancement of AI techniques and abundant cultural resources in Dunhuang area can promote the improvement of each other. This paper will discuss how to imply Artificial Intelligence techniques in Digital Humanities. It will also focus specifically on the applications of these techniques in Dunhuang culture studies.

2 Digital Humanities and Artificial Intelligence

2.1 Digital Humanities Development Direction

Humanities computing methodology has changed the academy of humanities in a subversive way, where it combines quantized orientation and qualitative research. Even before the emergence of the computer, people had tried to find the patterns hidden in mass data by analyzing, quantizing, and integrating them. After the computer was invented, the processing of mass data was promoted a lot thanks to the computing power and unbiased analysis of machines compared with humans.

At that time, although with the assistance of electronics, the essential methodology in the area of History Statistics was the same, which was integration, classification, and indexing of text. The development consisted in processing speed, which was raised to a new level with the strong ability of searching and organization of computer. This kind of innovation was called Humanities Computing.

Digital Humanities is the methodological community of Data-driven research of humanities, which is based on new technologies such as computers and networks (Liu & Ye, 2017). As was mentioned by Liang Qichao, a famous historical figure of China, a new research methodology called History Statistics, which applied statistics in history learning to recognize general historical trends and then to create new knowledge, should be implied in studying. “Those subtle changes in society should be collected and analyzed together to find out the most valuable principle.” Digital humanities are expected to achieve more than digital computing; nevertheless, humanities computing facilitated the development of Digital Humanities because people had been aware of the power of machines’ ability to process data. One aspect was the building of several libraries and archives centered on the Internet (Li, 2021).

In 2008, Stanford University described the planned but unrealized interdisciplinary program of digital humanities as follows: “The growing importance of digital technologies in contemporary culture has led to new forms of academic research, new ways of evaluating organizational humanistic knowledge, and new forms of cultural communication. The core concerns of the humanities have shifted. How will reading and writing change in the digital age? What new forms of cultural expression have emerged in the digital age? How do they build on or break the old form? How should we evaluate the moral and political implications of digital technology? What tools do we have or need to develop, understand, and/or leverage this new technology?”[23]
In Digital Humanities, data were put in general content. This required people-computer coworking, using already created digital data, quantized models, specific knowledge of a subject, and a combination of skills. This means the function of the computer should be far more than text processing. Up to now, as electronic technology’s developing direction is mainly toward Artificial Learning, Digital Humanities research is combined with artificial intelligence too.

### 2.2 the Relation between Digital Humanities and Artificial Intelligence

The “Methodological Commons” of digital humanities are summarized as discovering, collecting, comparing, delivering, and collaborating, which is realized by various technologies, such as digitization, data management, and analysis, visualization, VR/AR, and machine learning (McCarty & Short, 2017). The “Methodology Basis” of Digital Humanities is mapping and simulation, which refers to building a system in the virtual world using digitalized research and historical materials and realizing emulating real scenes, where it is necessary to build abstract concepts and logic structure that is simultaneous with the real world; therefore, the construction of digital humanities can be classified into 3 aspects, which are structuration and visualization, reconstruction and propagation, and computer-assisting development and extension (Luo, 2015).

With these methodology theories, we can conclude that the computer, as the basis of Digital Humanities research, provides requisite tools, platforms, and materials for it, and the computer can facilitate the integration and intersection of different humanities subjects.

### 2.3 the Deficiency of Machine Applying in Digital Humanities and Dunhuang Culture Studying

Compared with the Humanities Computing method, Digital Humanities has facilitated a new form of humanities research in the digital area. This new nomenclature explicitly means a larger field and a broader context to describe the activities and structures between the humanities and information technology as a whole[p]. Ulteriorly, in [q], Matthew and Lauren included new technologies’ applications such as AR games, 3D modeling of cultural relics, and label activism and its analysis. However, the applications of these newest technologies are mainly focused on subjects such as history, library science, literature, and linguistics (Huang & Li, 2020; Zhu & Nie, 2016). Although many researchers have studied the possibility of a combination between machine and intangible cultural heritage, the approaches are still looking at data collecting, processing, classification, and propagation (O, 2016; Zhang et al., 2021; Zhang & Wang, 2021). Some researchers have also discussed the most advanced technologies’ applications in intangible cultural heritage, such as VR and AR, but they hardly propose possible future methods to apply Artificial Intelligence in humanities. Another situation is that, although Dunhuang Culture is famous in China and abroad, few of its intangible cultural heritage has been included in World and National Intangible Cultural Heritage Lists; therefore, compared with that in some other areas, the intangible cultural heritage in Dunhuang have deficient attention, and few types of research have been done to apply newest technologies in them.

In this study, the application of Artificial Intelligence in humanities will be discussed in a systematic way, relating the 3 directions of future development of Artificial Intelligence with 3 aspects of intangible cultural heritage protecting, inheriting, and developing. Specifically, the study will pay much attention to Dunhuang’s intangible cultural heritages and introduce some approaches to applying the newest technologies.

### 3 Smart Data Driving Learning for Structuration and Visualization

#### 3.1 Smart Data Driving Learning Combined with Digitalization

**3.1.1 From Big Data to Smart Data**

The mainstream of Artificial Intelligence was Data Intelligence, where research mainly focused on the collection and integration of data; while in the future, combining Smart Data with Deep Learning is forecast to be prevalent, where the human being will be superseded by machines in data analysis and decision making (Zhang & Wang, 2021). Through effective processing, compared with merely Big Data, Smart Data is possible to provide us with actionable information and to improve decision making by finding the hidden patterns of mass data, which is hard to realize by human working (Zeng, 2017).

**3.1.2 Smart Data Intelligence in Humanities**

Besides data accessible on the internet, those paper-based data and real objects should be included as well. These data called unstructured data, are usually large in quantity and diverse in type, nature, and quality, which makes them hard for people to tag, classify, and process. The basis of smart-data learning is digitalization, where organization and integration of data are essential. Besides this, Artificial Intelligence facilitates the perception, inference, and innovation of machines. For example, the famous AlphaGo deep reinforcement learning first learned human experiences; then it built its own pattern of chess game; finally, in mass-repeating and high-speed learning, it concluded new methods of its own and applied them in the game with humans (Tian, 2016).

Structuration and visualization of digital humanities are facilitated by the digitalization of literature and data, for example, by collecting images and including paper-based materials in databases. Therefore, it is possible for researchers to analyze these
digitalized materials in a more systematic and comprehensive way, contributing to advanced analysis of the materials. From traditional digitalization technologies such as images, videos, and audio, to emerging technologies such as holography, 3-D scanning, and motion capture, digital technologies have multiple advantages in collecting, preserving, and recovering cultural heritage. For example, after digitalization, databases and search engines will be built for preserving and consulting in traditional Big Data Approaches (Huang, 2015). This way of utilizing literature is still in the original state of paper replacement, presenting only the sequential display of time, lacking spatial, sequential display (Yang, 2007). While for smart data technologies such as Data Mining and Visualized Analysis of Texts, besides retrieving the full text, they can also change the structure of the original data by adapting to their own need for research and reorganizing the original information to generate new knowledge and information automatically. Take special-visualized analysis as an example; in Geographic Information System (GIS), the mass, static, and scattered digitalized ancient literature in Chinese databases are integrated and displayed on maps on a large scale; here, the geographical distribution of relevant authors is presented on a map based on the thread of ancient books (Wang, 2011). This can assist researchers in summarizing the geographic distribution features of languages, histories, and cultures, utilizing diversified cultural values in the rich ancient literature of China.

3.2 Applications of Smart Data Driving in Dunhuang learning

In terms of Dunhuang Studies, there are three types of objects of digitalization, which are document literature, cultural relics, and other intangible cultures, such as Quzixi (a kind of opera prevent in the Dunhuang area).

3.2.1 Deep Reinforcement Learning and Natural Language Analysis

The analysis of natural language in Dunhuang Culture studying is important because Dunhuang culture is a civilization synthesis of different regions and eras. Cross-learning can greatly facilitate the new development of Dunhuang research, where researchers from different cultural backgrounds and subjects work together to study the same topics and complete the same projects. There is a problem that people have to read related materials first to have a general knowledge of the area, but the texts are mostly in Chinese. Therefore, translation is essential in the cross-learning of Dunhuang culture.

Traditional machine translation is improved by the neutral network and end-to-end neutral machine learning, so new translation methods can realize more accurate translation (Xi & Zhou, 2016). Compared to old machine translation techniques, it is more appropriate in text with sentiment and complex word meaning and syntax, which are the features of Dunhuang literature.

3.2.2 Computer Vision and Image Style Transfer

Image Style Transfer has been utilized widely in people’s life. For example, some image processing apps provide image style transfer services that people can mix together their photos with other photos and change the style of their own photos (Tang, 2018).

This technique can be implied in Dunhuang culture, whose great proportion is consisted of sculptures, murals, dances, and calligraphies. Visualization of these materials is an important part of Dunhuang’s studying. Dunhuang elements can be seen everywhere now, such as in movies and modern dances. Image style transfer can combine them together and create new cultural productions. For example, in cartoons, turning Dunhuang murals into cartoon styles can be done simply by putting them with cartoon pictures and using the image style transfer technique; in cultural and creative products, texts can be shown in the style of calligraphy using this technique as well.

4 Cross-media Intelligence for Reconstruction and Propagation

4.1 AR and VR

Chinese researchers described cross-media computing as closely mixing text, image, audio, and video (Pan, 2016). The manifestation of cross-media is AR and VR. Combined with other techniques such as motion capture, they can help to reconstruct and present the aiming object with high accuracy, immersive experience, and multiple dimensions.

Augmented Reality is mainly realized by technologies such as multimedia, 3D modeling, real-time display, and sensors (Chen, 2022). It integrates information in both the real and visual world, mapping visuality signals that can be perceived by people, such as sound and images, to the real world. AR can also make adaptations to visual information according to changes in the real world.

Visual Reality actually realizes an inverse process of AR, that it maps reality information to the visual world and builds a 3D environment in which people can have perceptions of visual signals as well.

AR and VR are applied in many scenes now, such as VR games, where people can wear goggles and gloves and do some actions to interact with the system and get related feedback. In the clothing industry, using VR and AR, designers can preview the 3D
effects of the design and make improvements to it until they are satisfied. After that, customers can try clothes by putting on clothes on screens to see if they fit.

4.2 Applications of VR and AR in Dunhuang Digital Museum

There are 2 patterns of the digital museum, which are web browsing and immersive interaction. The first one provides a searching service for the audience on the Internet, and people can enjoy the sight of exhibits in the digital museum from any angle they want. Additionally, these exhibits are shown in great detail so that people can even get a better experience looking at them on the Internet than in a real museum.

In Dunhuang, many of the caves are in bad conditions that people cannot go into there to visit. However, only a picture is not enough for people to get a full view of the scenes in the caves; here, the digital museum technique is very useful. The sculptures can be displayed in detail with 3D visual effects on the Internet; the murals can be shown as 3D images when people put on human-machine interface devices such as goggles; they can see the figures in the murals floating and dancing and even interact with the figures.

4.2.1 3D Exhibition of Sculptures

Painted sculpture 3D exhibition on the Internet is not easy to realize because the sculptures are of various and complex forms. 3D modeling and image-based rendering (IBR) are good methods to solve this problem. People can build visual 3D figures of sculptures with realistic effects, such as shadow, surface, and specification (Shi et al., 2002). By 3D scanning and taking photos, data can be collected and then processed; finally, people can depend on these data to create a 3D model with great detail that can be exhibited in digital museums. Moreover, these models can be used in AR displays. They are superimposed into the real scene to form interactive 3D images when people are wearing goggles or looking through electronic devices such as their phones, where people can look from different angles and get a full view of the sculptures.

4.2.2 VR Experience of Murals in Mogao Caves

The murals in Dunhuang Mogao caves are becoming more and more fragile under the threat of seepage, wind erosion, and sunniness (Fan, 2016). Therefore, how for the visitor to visit the caves without actually being there has become a problem. Only “looking” is not enough for people to get an immersive perception; to “really be there” is a good way to realize this kind of experience. Using dynamic modeling, real-time dynamic rendering, human-computer nature interaction, stereo display and sensing, and 3D virtual sound, people can get an immersive experience, like they are really visiting the caves and enjoying the arts. In the VR world, different elements can be added, such as texts, pictures, and videos, so while people are visiting the scenes in VR, they can get even more information about the caves and sculptures, including the craftsmanship, history, and preserving situation.

5 Human-machine Association for Renovation and Recreation

5.1 Machine and Human Intelligence Comparison and Connection

Computer-assisting development and extension is a good way to facilitate the renovation and recreation of humanities. New technologies are ideal tools for humanities research because they can create new environments of traditional culture that adapt to the new era, find more space for traditional culture, and partly resolve the conflict between modern civilization and cultural heritage. Artificial intelligence and machine learning, for example, as the mainstream of the technology revolution, has contributed to all aspects of the protection, propagation, and revolution of traditional culture.

Machine learning has many advantages compared with human learning, but its application is still limited because it is hard at present for the machine to fully imitate the behavior mechanism of people, leading to problems such as non-convergence (Cai & Song, 2015). This limitation is presented noticeably when it comes to digital humanities, where knowledge is mostly based on people’s philosophical contemplations, which are unstructured, immeasurable, inconsistent, flexible, and infinite. These bits of knowledge cannot be concluded into processable data to become the input for the machine, which can only accept certain physical stimuli, but not abstract concepts (Ma & Zou, 2017). One good solution for the application of machine learning is to add human intervention.

5.2 Applications of Human-machine Association in Dunhuang

5.2.1 Computer-Aided Murals Restoration

The restoration of murals of Dunhuang is usually done by experienced craftsmen, but sometimes there can still be their fault because the murals are unique, and they cannot practice under the same condition to see if this time they can succeed. Using the computer to simulate the real situation is a good way to reduce the chances of failure.

The human-machine circle in this process is essential. First, people will use software to preprocess the images to improve their quality; second, the machine will learn about the damage to different parts and get the restoring objects; third, people will input
some information about the restored mural, such as dynasty, cave, content, style, and painting techniques, and the computer will
detect which kind of damage (discoloration, fading, or exfoliation) is of the mural; forth, depending on the information, the machine
will choose from the restoring techniques and search for experiential and analogous knowledge; fifth, the machine will adapt these
bits of knowledge and techniques to the targeting mural and fill color and create patterns; finally, machine and human will work
together to merge all the restoring objects to the original mural in sequence (Lu et al., 2002). The result is shown visually on screens,
and people can try to adapt the result. Finally, the experienced craftsmen can redo this on the real mural.

6 Discussion
We discussed how people could combine humanities research with computer technologies, especially today, and how to improve
the research of digital humanities by combining it with Artificial Intelligence technologies. Specifically, these technologies can be
implied in the studies of Dunhuang Culture because it’s a multi-background area and needs to be explored by researchers from
different civilizations and areas.

The traditional Digital Humanities research focused mainly on text classification, analysis, and indexing, while the new one
concludes with techniques facilitated by new technologies. The new direction of Artificial Intelligence development includes Smart
Data Intelligence, Cross-media Intelligence, and Human-machine Association Intelligence, which are appropriate for the 3
important aspects of humanities research, especially intangible cultural heritage. The 3 aspects are listed as Structuration and
Visualization, Reconstruction and Propagation, and Renovation and Recreation. We discussed the methodology to realize them
with the latest AI technologies and the application in Dunhuang Culture Research.

While the contribution of AI techniques in Dunhuang studies is great, it brings some problems. The first is, how do we define the
production of machine intelligence? Where do we put cultural heritages when they can be done by machine? Does volume
production diminish the value of intangible cultural heritages? For those who love traditional culture, it might be a paradox that
the essence of intangible heritages lies in history and humanities meaning of them, and the reason for them to be so precious is
that they show the temperature of human nature, wisdom, and contemplation; if produced by machine, where should we go to
find the meaning of heritages?

Another problem is related to the propagation of traditional culture. When new technologies promote the propagation of cultural
heritages, people are worried about inappropriate changes brought about. Propagation and innovation are necessary and
important in culture inheritance, but will they change the nature of the original culture? Will the culture be too affected by other
cultures and lose its features, becoming homogeneous with others?

7 Conclusion
Digital Humanities and Artificial Intelligence are forming a new scene of humanities research. Besides text data, other media can
be digitalized, processed, and analyzed by machines now. Digitalization is the basis of digital humanities, which was realized before
the invention of Artificial Intelligence, but people were looking for more possibilities from the machine in humanities. This
intersection not only facilitates the development of humanities but also greatly increases the speed of knowledge production.
Therefore, merely using a machine to analyze data for human learning is not enough now. More advanced technologies should
be implied to promote knowledge fusion and propagation. The digital humanities studying process is concluded as discovering,
collecting, comparing, delivering, and collaborating; the latest AI technologies include Smart Data, Cross-media Intelligence, and
Human-machine Association. The latter is just appropriate for promoting the former. In the article, we specifically talked about the
application of Dunhuang culture learning because this culture is typical and suitable for the research of interdisciplinary and
intangible cultural heritage restoration and renaissance. Dunhuang is famous for its painted-sculptures, murals, and other cultural
relics; it is also famous for rich historical literature and cultural skills. In the article, we discussed the techniques to utilize Artificial
Intelligence technologies in Dunhuang culture’s inheritance, development, and innovation. With Smart Data combined with
structuration and visualization, people can extract more valuable information by looking into the material with a comprehensive
view assisted by computer processing. These materials can be visualized and combined with time and space information, giving a
clear thread for researchers to find out the hidden information that is not easy to be concluded through pure texts. Cross-media
Intelligence combined with reconstruction and propagation, using VR and AR techniques, visitors can get an immersive experience
looking at and interacting with the real scene and objects such as the sculptures in the caves. This kind of visit makes the journey
more fun and makes cultural relics more lively, making it easier to arouse the interest of people to learn more about the intangible
cultural heritages of Dunhuang. Human-machine Association combined with renovation and recreation, the craftsmen that are not
skilful with new technologies will have access to using the computer to restore precious murals with fewer errors while taking less
time and effort. In the future, with the development of Artificial Intelligence and with more research into digital humanities, the
application can be extended more. While there are developments, there will also so some new emerging problems. One that
should be paid attention to is the positioning of machine-creating cultural items. Cultural products are supposed to be closely
related to the feelings and creativity of humans, so those created by machines can be counted as cultural products is an issue worth discussing.

**Funding:** Please add: This research received no external funding

**Acknowledgment:** This paper is an output of the project of Desktop Research on Dunhuang Intangible Cultural Heritage. My sincere thanks and appreciation go to the sponsor and the keynote speaker Gu Changwei.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


The Application of Artificial Intelligence Technologies in Digital Humanities: Applying to Dunhuang Culture Inheritance, Development, and Innovation


