AI-Enabled Cultural Heritage Conservation Data Management: Taking Mogao Grottoes as an Example

Shunren Wang [1,2], Yipu Gong [2,*], Xiaowei Wang [2], Kui Jin [2]

[1] Gansu Mogao Grottoes Cultural Heritage Conservation and Design Consulting Co., Ltd, Dunhuang, Gansu, China - wangsr @dha.ac.cn
[2] Dunhuang Academy, Dunhuang, Gansu, China - (gongyp, wangxw, jinkui) @dha.ac.cn

Keywords: Cultural Heritage Data Management, Mogao Grottoes Wall Paintings, Condition Assessment Data, Graph Data Structure, AI Technology.

Abstract

This paper explores the application of AI technology in cultural heritage data management, focusing on wall paintings’ condition assessment data from Mogao Grottoes, and constructs a framework integrating graph data structures with an AI model. The method integrates multi-source data from Mogao Grottoes wall paintings surveys, such as handwritten records and digital archives, to facilitate efficient analysis and rapid query of deterioration and spatio-temporal information. Leveraging this novel technical framework, the study enhances the intelligence of cultural heritage data management, offering valuable approaches for the conservation of similar heritage sites. The findings effectively advance the digital and intelligent transformation of cultural heritage conservation, aligning with the focus on data-driven diagnosis for conservation decision-making.

1. Introduction

As a world-renowned cultural heritage site, the Mogao Grottoes represent the largest and best-preserved treasure trove of Buddhist art globally, boasting immense historical, artistic, and scientific value (Fan, 2000). The Mogao Grottoes have been preserved for thousands of years, which is closely related to its occurrence environment (Li et al., 2010). Affected by their own muddy materials and natural environmental factors, wall paintings that have survived through thousands of years have developed various damages such as flaking, efflorescence, hollowing, and fading (Chen, 2017). These issues not only compromise the artistic integrity of wall paintings (Pei and He, 2020), but also threaten their long-term survival (Li et al., 2013). To address this, we have established an early-warning monitoring platform to continuously track microenvironmental parameters (temperature, humidity, CO_2 concentration) within caves and their surroundings. This enables proactive mitigation of deterioration progression through precise environmental control, complemented by regular visual inspections to monitor dynamic changes in deterioration. Decades of wall paintings conservation efforts have yielded a vast corpus of records (both handwritten and digital), documenting wall paintings content, deterioration types, spatial distribution, developmental changes, and historical conservation interventions. These materials form the basis for long-term cave preservation and health assessment. However, the complex structure of the Mogao Grottoes, where each cave wall harbors multiple deteriorations, poses challenges for data analysis. For instance, tracking deterioration trends across caves over time involves massive datasets, labor-intensive documentation review, and difficulty in comparative analysis. Traditional relational database management exacerbates these issues; it requires writing lengthy multi-table join queries (Gong et al., 2025), and incurs substantial data collection delays. By relying on tabular organization, traditional systems struggle to flexibly represent the spatio-temporal relationships among caves, walls, and deterioration, hindering data management and impeding evidence-based conservation decision-making.

In contrast, graph-structured data using nodes and edges intuitively connects caves, walls, and deterioration. Each cave can be modeled as a node, with wall and deterioration locations as sub-nodes, while edges define spatio-temporal relationships. This framework enables rapid deterioration localization and detailed spatio-temporal evolution mapping, facilitating dynamic deterioration monitoring.

Recent advancements in AI technology, from theoretical research to cross-disciplinary applications (Allegra et al., 2022), have transformed fields like image recognition (Cheng et al., 2018), where deep learning models excel at feature identification (Dubois et al., 2024), and natural language processing, where pre-trained models enhance translation and question-answering efficiency (Asmaa et al., 2023). In cultural heritage, AI has enabled museum intelligent navigation (Menotti, 2025; Wang et al., 2025), and data-driven deterioration diagnosis for targeted restoration planning (Yi et al., 2024).

Grounding our work in daily Mogao Grottoes conservation practices, we propose integrating graph data structures with AI models to develop an innovative wall paintings condition assessment system. The graph structure efficiently organizes multi-source data into an intuitive knowledge graph for rapid retrieval, while AI leverages its analytical power to mine graph data, supporting scientific conservation planning. This research not only elevates wall paintings conservation management at Mogao Grottoes but also provides a replicable framework for global heritage sites, advancing the digital and intelligent transformation of cultural heritage conservation.

2. Methods

2.1 AI-Assisted Structured Data Processing

In the long-term conservation surveys of Mogao Grottoes wall paintings, the earliest paper archives documenting wall paintings damage date to 1976 and systematic electronic record-keeping has been conducted since 2007. These materials meticulously document wall paintings deterioration across 492 caves, capturing protection status from multi-dimensional and temporal perspectives. For this study, 20 representative caves (4.1% of the total) were selected as experimental cases to validate the accuracy of the proposed method. Given the extensive scale of the 492-cave dataset, these 20 caves were purposefully chosen to demonstrate the method's effectiveness in handling diverse deterioration types (e.g., flaking, salt efflorescence, detachment), ensuring the method's validity can be rigorously tested without overwhelming data complexity.

To efficiently extract information from handwritten records, we employ AI-OCR (Optical Character Recognition) tools. Historical handwritten materials with black carbon ink, remain remarkably legible due to the ink's durability. However, the idiosyncratic writing habits of recorders, marked by conjoined strokes and penmanship characteristics, pose challenges for automated recognition. Despite these constraints, the system has achieved an 85% overall accuracy rate in extracting key information from handwritten materials, as validated by balancing ink clarity with specialized processing for handwritten nuances (Figure 1). All identified errors are subsequently manually corrected by heritage conservators to further refine data accuracy.

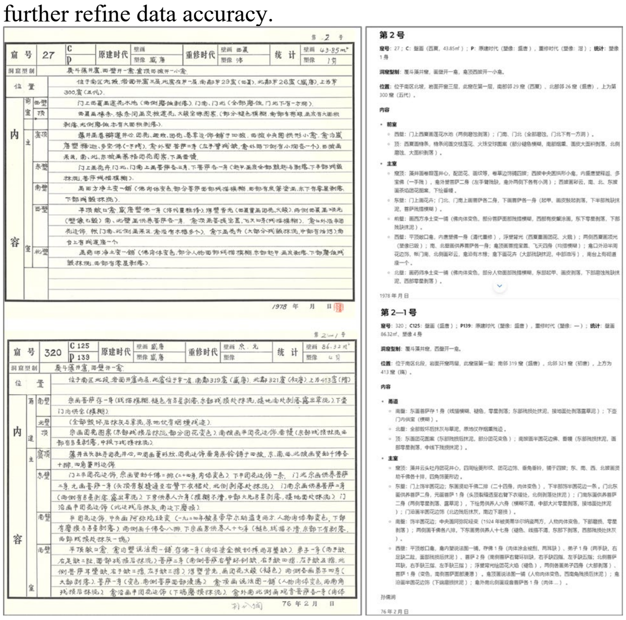


Figure 1. OCR processing example of handwritten Records.

To efficiently structure semi-structured data, we adopt a human-supervised, rule-based AI-assisted approach, where each electronic document is manually processed to guide the AI in entity extraction. Leveraging deep learning-based models (e.g., Doubao, DeepSeek) trained on extensive corpora and predefined rules for cave numbers, deterioration types, and spatial locations, the AI generates entity annotations, which are then meticulously reviewed by heritage experts to ensure accuracy. This hybrid workflow combining AI's contextual understanding with rule-based parsing and human oversight-enables us to extract key information such as cave number, deterioration location, type, and discovery time from each document individually. For deterioration identification, we pre-set comprehensive keyword rules via AI, covering flaking, salt efflorescence, and other deterioration types. The system then intelligently extracts quantitative severity data (area, length), location, descriptions, and repair status from complex texts, with each deterioration in the same location analyzed independently.

Following the AI-driven extraction, all validated entities are manually saved as standalone JSON files centered on cave numbers. These JSON files are then converted to Excel format to standardize the dataset, facilitating seamless import into graph databases for further analysis (Figure 2). This process ensures manual supervision of each step from AI generation to format conversion, balancing algorithmic efficiency with human precision in documenting wall painting deterioration across six dimensions: type, severity metrics, location, description, emergency repair necessity, and repair history.

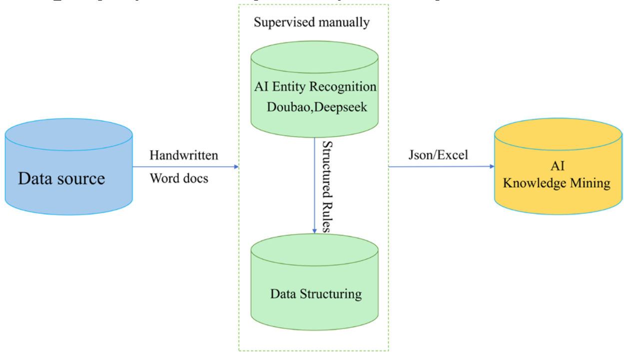


Figure 2. AI-structured process of wall paintings protection data in Mogao grottoes.

2.2 Cave Spatial Relationship Model Based on Graph Data Structure

Mogao Grottoes comprise 492 caves distributed across different levels, with some caves adjacent to each other even sharing walls where wall paintings exist on both sides. Traditional relational databases face significant limitations in describing this complex spatial structure. Relational databases store data in tables, each associated with primary and foreign keys, but the spatial relationships of Mogao Grottoes defy simple linear or hierarchical modelling. For instance, cave locations include anterior chambers, main chambers, Yongdao, etc., each subdivided into east, west, north, south walls, ceilings, niches, and more. The shared walls between caves, e.g., the south wall of one cave's main chamber being the north wall of another cave's main chamber is difficult to represent concisely in relational databases. Forcing such relationships into a relational model results in overly complex table structures, inefficient join queries, and an inability to intuitively reflect spatial dynamics.

Graph data structures effectively address these challenges by integrating each cave's unique spatial architecture with commonalities in wall painting deterioration. Specifically, the model defines four core node types: Cave nodes (identified by unique Cave_number), Cave location nodes, deterioration location nodes and wall painting deterioration type nodes, which predefine categories such as flaking, powdering, and paint loss. Spatial relationships are formalized through directed edges: edges labelled adjacent connect Deterioration_location nodes when caves share walls (establishing cross-cave spatial links between corresponding deterioration sites); Hierarchical containment is implemented via contain edges, forming a nested structure: Cave→Cave_ location→Deterioration_location; Each Deterioration location node is further linked to Deterioration types through With_wall_paintings_Deteriorations edges, which record attributes like deterioration area (e.g., 0.8 m^2), repair status (Is_it_fixed), and quantification data. This node-edge modelling clarifies the spatio-temporal interplay between Mogao Grottoes' structural space and wall paintings deterioration information, supporting dynamic analysis of multiple documented deterioration types. The framework's ability to represent both containment hierarchies and lateral adjacencies is visualized (Figure 3), laying a robust foundation for data-driven conservation decision-making.

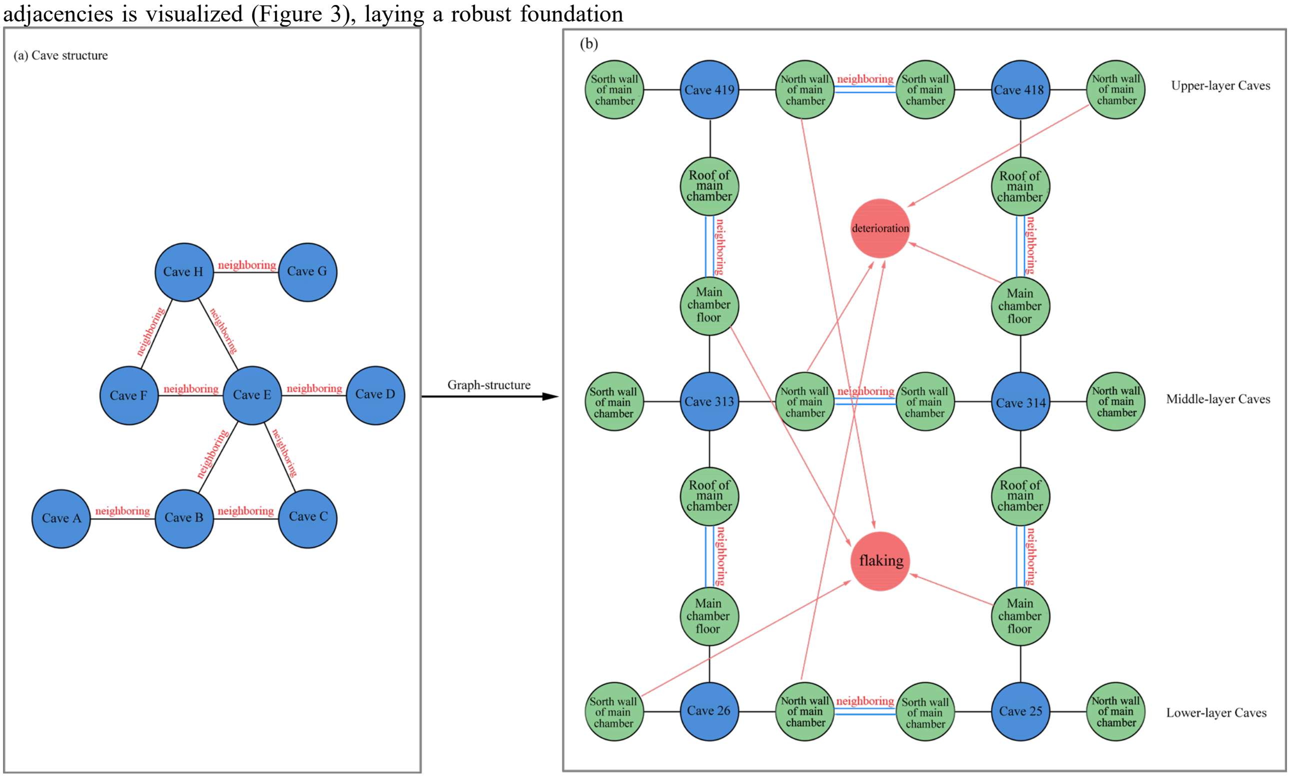


Figure 3. Schematic diagram of graph-structured data for caves spatial layout in Mogao Grottoes.

2.3 Retrieval of Wall paintings Deterioration Map Structure Database by AI

With the continuous advancement of systematic condition assessment for Mogao Grottoes wall paintings, accumulated deterioration data exhibits characteristics of massive scale, complex structure, and dynamic growth. Traditional manual retrieval faces critical bottlenecks: conservators spend hours reviewing records, hindering emergency response, while manual analysis is prone to subjective biases: fatigue and professional discrepancies often cause keyword extraction errors.

Although traditional relational databases enable basic keyword matching, they struggle with complex semantics and multi-dimensional logical relationships. Take the query “Find records of flaking-related deterioration added after 2020 with an area > 0.3 m² on the main room’s east wall”–traditional systems rely on single-keyword matching, failing to analyze logical relationships among spatial location (main room east wall), area threshold (0.3 m²), time range (post-2020), and deterioration type (flaking). In practice, such searches may incorrectly return “flaking < 0.3 m²” or “pre-2020 flaking-like issues” due to lacking semantic understanding, causing result deviations from real needs. This inefficient retrieval mode has become a core technical barrier for Mogao wall paintings conservation. In critical scenarios like relic deterioration emergency response or preventive conservation planning, the inability to obtain timely accurate data undermines the scientific basis for conservation decisions, delaying research and risking missed intervention opportunities. Thus, introducing AI-powered intelligent retrieval and constructing an efficient data system are urgent to shift Mogao conservation from experience-driven to data-driven.

Graph database technology has emerged as a cornerstone for cultural heritage data management, leveraging its core advantage in processing complex relational data. The full-process system herein adheres to strict verification: first, heritage experts perform multi-round manual validation on raw data to ensure accuracy (e.g., deterioration location, type); next, a hierarchical data structure is automatically constructed in Excel via preset templates, using a standardized field system Cave-Cave_Location-Deterioration_Location-Deterioration_Type-Quantitative_DataDescription-Repair_Status for structuring; finally, the validated dataset is imported into a Neo4j graph database to form a semantically associated knowledge graph, laying a structured foundation for intelligent retrieval.

The AI model uses deep learning to parse semantic association relations between graph nodes, upgrading traditional keyword matching to implicit relationship mining, which is supported by pre-validated Excel data ensuring graph database semantic consistency. For dynamic data updates, the AI leverages Neo4j’s incremental learning framework for real-time adaptation, while Excel’s pre-structured templates enable cross-source verification with graph data. Regarding hallucination control, the system integrates a dual manual verification mechanism: conservators first perform secondary data validation in Excel, then subsequent reviews of critical AI outputs (e.g., emergency repair recommendations), correcting deviations by cross-referencing original records. This triple guarantee rigorous preprocessing graph semantic reasoning—continuous human validation—ensures retrieval reliability and minimizes inference errors, fully addressing Mogao conservation’s dynamic decision needs.

When users pose natural-language queries, the AI model parses questions to extract key elements and their relationships, then transforms them into graph-structure query statements per preset rules. Based on these, the model locates corresponding nodes in the graph database, filters qualified results, sorts them by relevance, and prioritizes presenting the most valuable findings to conservators, enabling quick and accurate information acquisition (Figure 4).

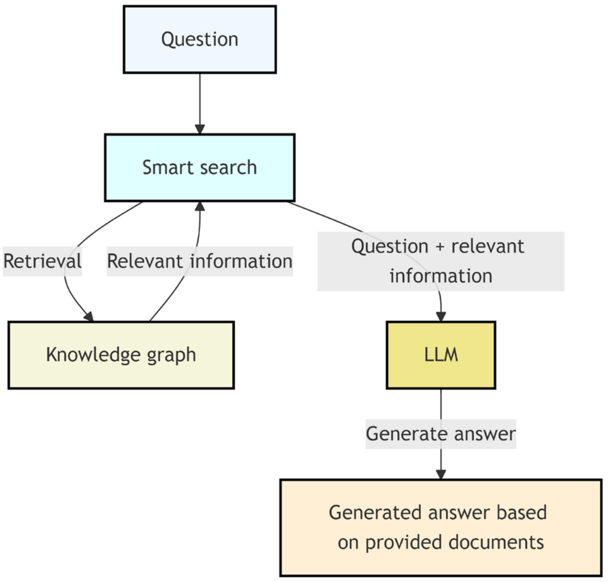


Figure 4. AI realizes the retrieval of Wall paintings deterioration graph-structured database.

Based on the above technical roadmap, we have developed a novel data management approach for Mogao Grottoes wall paintings conservation decisions. The core research question is whether this approach can be effective in practice and address the conservation decision-making challenges faced by wall paintings preservation. Subsequent validation will integrate practical cases and data, starting with graph database-based spatial structure modeling of grottoes, conducting cave correlation analysis via deterioration characteristics, and applying AI-graph data integration for associative analysis. These steps will verify the new approach's effectiveness and feasibility while exploring its potential in wall paintings conservation. This not only provides a scientific basis for subsequent conservation planning decisions but also enhances the overall level of Mogao wall paintings conservation, driving it toward more efficient and scientific development.

3. Results

3.1 Constructing the Spatial Structure Relationships of Caves with a Graph Database

We adopted Neo4j as the platform to construct graph-structured data. Neo4j enables high-performance processing of complex relationships, and its Cypher query language streamlines data operations, effectively representing cave-related data to facilitate wall paintings deterioration analysis. In the conservation of cave wall paintings, constructing grotto spatial structures via graph databases is crucial for accurately assessing deterioration status. The cave space is partitioned, with different wall paintings–bearing areas modeled as nodes in the graph database. Each node records detailed attribute information: deterioration type, location, severity, progression trend, and repair history of the corresponding area.

Spatial relationships between deterioration areas (e.g., similar progression and severity in adjacent caves, or potential transmission due to structural interconnections) are represented by edges in the graph database. Edges also reflect whether deterioration damage in lower-level cave wall paintings correlates with groundwater rise and migration. With the constructed graph structure, spatial correlations of cave wall paintings deterioration are readily apparent. Researchers can leverage the graph database's query and analysis functions to quickly identify inter-deterioration relationships. We selected Cave 26 and its adjacent areas of Mogao Grottoes to construct the graph structure (Figure 5).

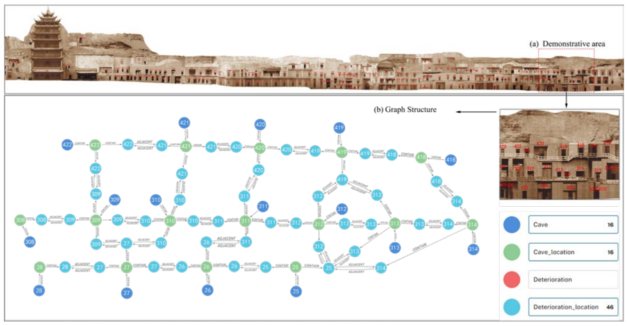


Figure 5. Cave spatial structure graph data schematic.

3.2 Cave Correlation Analysis Based on Deterioration Characteristics

Leveraging the established graph database, we can deeply explore potential connections between cave deteriorations. Through accurate deterioration data retrieval, close links between deterioration in adjacent caves become discernible. Moisture is a primary factor driving wall paintings deterioration in Mogao Grottoes. For instance, efflorescence and disintegration damage were identified on the north and south walls of Cave 26's main chamber.

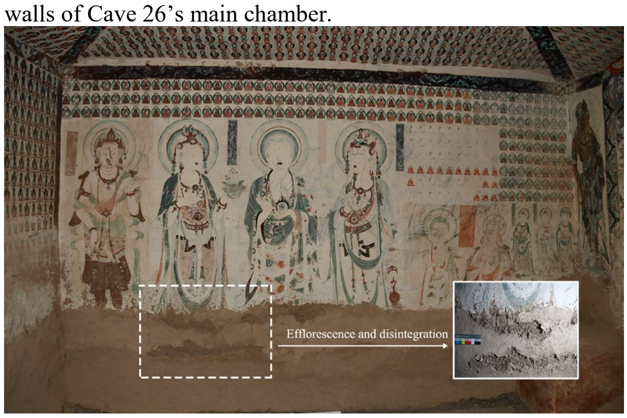


Figure 6. Efflorescence and disintegration on the south wall of the main chamber of cave 26 of Mogao grottoes.

When querying the graph database, the system swiftly locks onto cave nodes with recorded deterioration. By exploiting spatial relationships encoded in node edges, we efficiently trace adjacent cave nodes in the graph. Analyzing deterioration attributes of these nodes helps identify whether similar causes exist or if issues arise from spatial conduction and structural interactions. Taking Cave 26 as a case study, we searched for deterioration associated with adjacent caves. Graph database association analysis revealed that three adjacent caves (25, 27, 28) exhibited salt efflorescence and disintegration. Further inspection of cave wall deterioration distribution showed Cave 26 has 4 main chamber walls, 3 anterior chamber walls, and 1 Yongdao wall. Notably, Cave 27—adjacent to Cave 26's main chamber south wall—lacks north wall deterioration in its main chamber. This suggests Cave 26 experiences greater humidity fluctuations and water accumulation, likely due to being an open cave (affected by tourist-exhaled moisture), whereas Cave 27 is a closed cave with a stable environment (Figure 7).

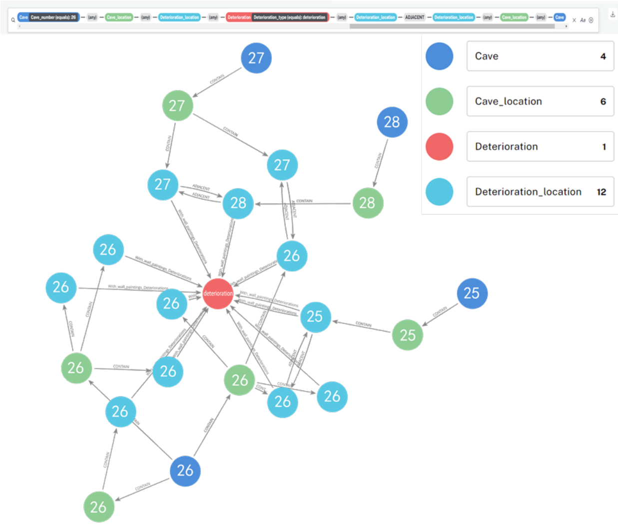


Figure 7. Search for deterioration in Cave 26 through the graph database.

Mogao Grottoes also presents special cases: during rainy days, upper cliff cave wall paintings suffer water seepage damage. The graph database can search for similar water-induced deterioration in surrounding caves. By comparing deterioration attributes of adjacent nodes, we determine if deterioration has spread. For example, if multiple adjacent caves show varying degrees of ceiling moisture or cliff detachment—spatially aligned with a rainwater leakage point—it indicates the seepage area affects surrounding caves, necessitating timely regional conservation measures to prevent deterioration spread. This graph database–based deterioration correlation analysis provides a comprehensive decision-making foundation for cave wall paintings conservation, enabling efficient identification of deterioration transmission paths and proactive risk warning for adjacent caves.

3.3 Application of Association Analysis Combining AI with Graph Data

In the research of wall paintings conservation in Mogao Grottoes, the combination of AI and graph database provides key support for wall paintings deterioration research and conservation decision-making. Graph database organizes data via nodes and edges, intuitively associating various information and defining spatio-temporal relationships through edge attributes. Natural language queries enable users to convert queries into executable graph database statements, ensuring quick and accurate information retrieval. Compared with the traditional approach of writing long, complex query statements, they offer obvious advantages. This combination also facilitates deep mining of latent information behind the data.

The “Cave Wall Painting Deterioration Assessment System” we designed is a comprehensive tool for evaluating the deterioration status of cave wall paintings. The system features a natural language query panel allowing users to input questions; upon submission, the system processes the query and generates corresponding Cypher statements, which are displayed in the Cypher query panel. A graph visualization window intuitively presents the data structure, while a node statistics panel provides key metrics such as the total number of caves, types of deterioration, total number of data records, and monitoring time range. The query results panel shows detailed data outputs and includes a download function for saving results. In addition, the system incorporates an ECharts-based deterioration trend chart module, which, based on query results, tracks changes in wall painting conditions over time by deterioration type (Figure 8).

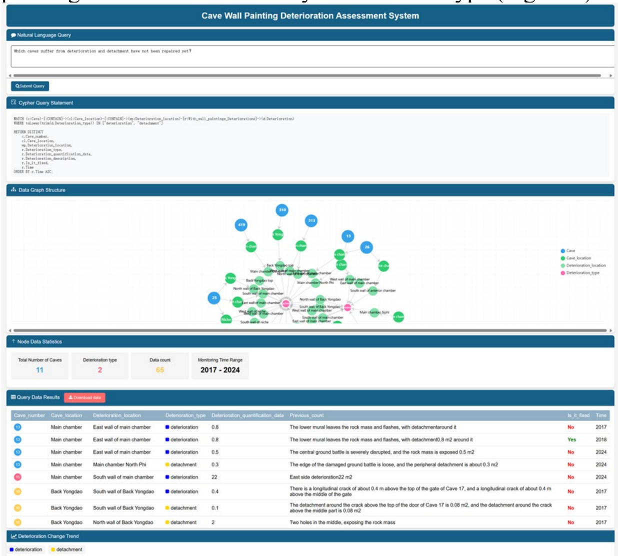


Figure 8. Schematic diagram of the functional module layout of the cave wall painting deterioration assessment system.

For instance, when identifying caves with deterioration in expansion, the AI system first mines graph - structured data. It then parses time - series statistics (covering attributes like type and quantity) of deterioration within individual caves, enabling accurate counting of each deterioration type’s occurrences from 2010 to 2024. Through this analysis, it identifies that 7 deterioration types across 4 caves (Nos. 13, 15, 16, and 85) show increasing trends (Figure 9). Cave 13 has the fewest types of deterioration, with only flaking showing a growth trend from 2019 to 2024. In contrast, Cave 15 has 5 types of deterioration, including Fissure, Mildew, Shedding, deterioration, and paint loss, all of which are on an upward trend, especially paint loss, which rises sharply after 2020. Cave 16 has four types of deterioration, all demonstrating an increasing pattern from 2021 to 2024. For Cave 85, from 2010 to 2011, both types (deterioration and flaking) exhibit a rising trend, then experience a decline, and maintain relatively stable levels afterward. Such fluctuations reflect the dynamic nature of the deterioration process and highlight the need for continuous monitoring to capture these changes accurately. This granular data visualization tracks deterioration records across types and caves over time, demonstrating the precision of AI - driven long - term heritage condition monitoring.

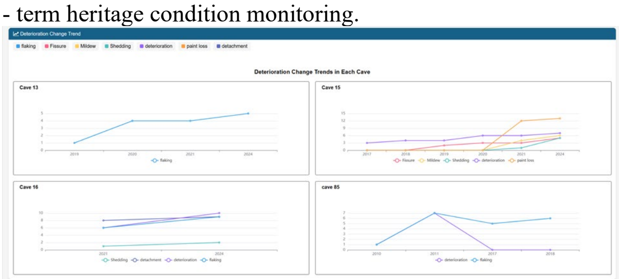


Figure 9. Retrieval results of increased cave deterioration.

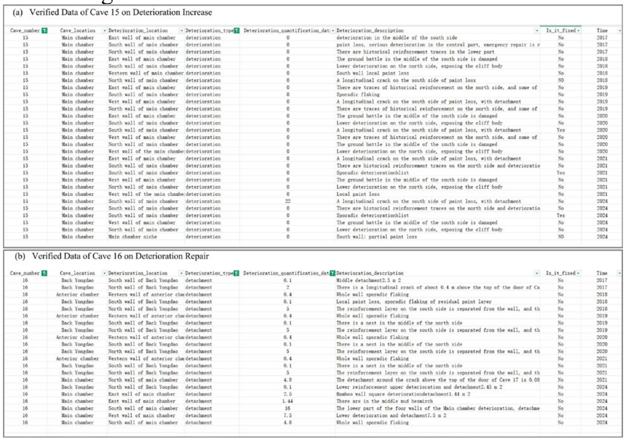
Similarly, when querying for caves with wall paintings deterioration and detachment that haven’t been repaired, the natural language processing module analyzes the query and converts it into a graph database statement. The AI model then performs a precise search in the graph-structured database, swiftly identifying and retrieving relevant nodes (e.g., cave numbers, deterioration locations) and edges (e.g., spatial adjacency relationships). Results are ranked by relevance, enabling efficient information extraction. The AI system rapidly retrieves structured data, including cave number, deterioration type, location, condition descriptions, and repair status, and the visualization uses color coding (green for repaired, red for unrepaired) to distinguish status. The system systematically quantified the area changes and repair statistics of two types (e.g., deterioration and detachment) in 11 caves from 2017 to 2024 (Figure 8).

Figure 10. Retrieval results of caves with deterioration and detachment without restoration.

The statistical results reveal distinct repair progress across different caves (Figure 10). For Cave 13, the total deterioration area stands at 2.40 m², with 1.60 m² remaining unrepaired, indicating that targeted restoration efforts have effectively reduced the damaged area. In contrast, Cave 16 presents a more pressing scenario: the total area affected by deterioration and detachment reaches 76.68 m², yet only 2 m² has been repaired to date. This significant gap between the extent of damage and completed restoration highlights the urgency of prioritizing this cave in future conservation plans. Such granular data on repair status and area changes not only clarifies the current state of each cave’s preservation but also provides concrete evidence for allocating resources, scheduling restoration tasks, and evaluating the efficiency of past interventions, thereby strengthening data-driven decision-making in long-term conservation strategies.

3.4 Data Validation and Reliability Analysis

To validate the accuracy of the system’s analysis results, we compared manually verified Excel baseline data with the system’s output (Figure 11). In the validation of 1,311 deterioration records from 20 sampled caves, the statistics on deterioration development trends showed that manual comparison of deterioration records across different years for each cave took approximately 3 hours, while the system completed full-volume data calculation in only 3-5 seconds. Taking the data of Cave 15 as an example, the statistical results of its deterioration types were completely consistent with the system output. In the analysis of deterioration repair status, manual counting of quantities alone required 0.7 hours, and if area change analysis was involved, it took 10 hours; however, the system generated a multi-dimensional statistical report including area fluctuations within 5-10 seconds. Taking the data of Cave 16 as an example, the area evolution data of its deterioration fully matched the manual verification results. Experimental data indicate that the system has met the expected objectives in both data processing efficiency and statistical accuracy, achieving an efficiency improvement of over 3 orders of magnitude compared to traditional manual analysis, and achieving an intelligent leap in cultural heritage conservation data management.



(a) Verified Data of Cave 15 on Deterioration Increase

Cave_number	Cave_location	Deterioration_location	Deterioration_type	Deterioration_quantification_dat	Deterioration_description	Is_it_fixed	Time
15	Main chamber	East wall of main chamber	deterioration	0	deterioration in the middle of the south side	No	2017
15	Main chamber	South wall of main chamber	deterioration	0	paint loss, serious deterioration in the central part, emergency repair is r	No	2017
15	Main chamber	North wall of main chamber	deterioration	0	There are historical reinforcement traces in the lower part	No	2017
15	Main chamber	East wall of main chamber	deterioration	0	The ground battle in the middle of the south side is damaged	No	2018
15	Main chamber	South wall of main chamber	deterioration	0	Lower deterioration on the north side, exposing the cliff body	No	2018
15	Main chamber	Western wall of main chamber	deterioration	0	South wall local paint loss	No	2018
15	Main chamber	North wall of main chamber	deterioration	0	A longitudinal crack on the south side of paint loss	NO	2018
15	Main chamber	East wall of main chamber	deterioration	0	There are traces of historical reinforcement on the north side, and some of	No	2019
15	Main chamber	South wall of main chamber	deterioration	0	Sporadic flaking	No	2019
15	Main chamber	West wall of main chamber	deterioration	0	A longitudinal crack on the south side of paint loss, with detachment	No	2019
15	Main chamber	North wall of main chamber	deterioration	0	There are traces of historical reinforcement on the north side, and some of	No	2019
15	Main chamber	East wall of main chamber	deterioration	0	The ground battle in the middle of the south side is damaged	No	2020
15	Main chamber	South wall of main chamber	deterioration	0	Lower deterioration on the north side, exposing the cliff body	No	2020
15	Main chamber	South wall of main chamber	deterioration	0	A longitudinal crack on the south side of paint loss, with detachment	Yes	2020
15	Main chamber	West wall of main chamber	deterioration	0	There are traces of historical reinforcement on the north side, and some of	No	2020
15	Main chamber	North wall of main chamber	deterioration	0	The ground battle in the middle of the south side is damaged	No	2020
15	Main chamber	West wall of the main chambe	deterioration	0	Lower deterioration on the north side, exposing the cliff body	No	2020
15	Main chamber	East wall of main chamber	deterioration	0	A longitudinal crack on the south side of paint loss, with detachment	No	2021
15	Main chamber	South wall of main chamber	deterioration	0	There are historical reinforcement traces on the north side and deterioratio	No	2021
15	Main chamber	South wall of main chamber	deterioration	0	Sporadic deteriorationblist	Yes	2021
15	Main chamber	West wall of main chamber	deterioration	0	The ground battle in the middle of the south side is damaged	No	2021
15	Main chamber	North wall of main chamber	deterioration	0	Lower deterioration on the north side, exposing the cliff body	No	2021
15	Main chamber	West wall of the main chambe	deterioration	0	Local paint loss	No	2021
15	Main chamber	South wall of main chamber	deterioration	22	A longitudinal crack on the south side of paint loss, with detachment	No	2024
15	Main chamber	South wall of main chamber	deterioration	0	There are historical reinforcement traces on the north side and deterioratio	No	2024
15	Main chamber	South wall of main chamber	deterioration	0	Sporadic deteriorationblist	Yes	2024
15	Main chamber	West wall of main chamber	deterioration	0	The ground battle in the middle of the south side is damaged	No	2024
15	Main chamber	North wall of main chamber	deterioration	0	Lower deterioration on the north side, exposing the cliff body	No	2024
15	Main chamber	Main chamber niche	deterioration	0	South wall: partial paint loss	NO	2024

(b) Verified Data of Cave 16 on Deterioration Repair

Cave_number	Cave_location	Deterioration_location	Deterioration_type	Deterioration_quantification_dat	Deterioration_description	Is_it_fixed	Time
16	Back Yongdao	South wall of Back Yongdao	detachment	0.1	Middle detachment2.5 m 2	No	2017
16	Back Yongdao	North wall of Back Yongdao	detachment	2	There is a longitudinal crack of about 0.4 m above the top of the door of Ca	No	2017
16	Anterior chamber	Western wall of anterior cha	detachment	0.4	Whole wall sporadic flaking	No	2018
16	Back Yongdao	South wall of Back Yongdao	detachment	0.1	Local paint loss, sporadic flaking of residual paint layer	No	2018
16	Back Yongdao	North wall of Back Yongdao	detachment	5	The reinforcement layer on the south side is separated from the wall, and th	No	2018
16	Anterior chamber	Western wall of anterior cha	detachment	0.4	Whole wall sporadic flaking	No	2019
16	Back Yongdao	South wall of Back Yongdao	detachment	0.1	There is a nest in the middle of the north side	No	2019
16	Back Yongdao	North wall of Back Yongdao	detachment	5	The reinforcement layer on the south side is separated from the wall, and th	No	2019
16	Anterior chamber	Western wall of anterior cha	detachment	0.4	Whole wall sporadic flaking	No	2020
16	Back Yongdao	South wall of Back Yongdao	detachment	0.1	There is a nest in the middle of the north side	No	2020
16	Back Yongdao	North wall of Back Yongdao	detachment	5	The reinforcement layer on the south side is separated from the wall, and th	No	2020
16	Anterior chamber	Western wall of anterior cha	detachment	0.4	Whole wall sporadic flaking	No	2021
16	Back Yongdao	South wall of Back Yongdao	detachment	0.1	There is a nest in the middle of the north side	No	2021
16	Back Yongdao	North wall of Back Yongdao	detachment	5	The reinforcement layer on the south side is separated from the wall, and th	No	2021
16	Main chamber	North wall of main chamber	detachment	4.8	The detachment around the crack above the top of the door of Cave 17 is 0.08	No	2021
16	Back Yongdao	North wall of Back Yongdao	detachment	0.1	Lower reinforcement upper deterioration and detachment2.45 m 2	No	2024
16	Main chamber	East wall of main chamber	detachment	2.5	Bamboo wall square deteriorationdetachment1.44 m 2	No	2024
16	Main chamber	East wall of main chamber	detachment	1.44	There are in the middle mud besmirch	No	2024
16	Main chamber	South wall of main chamber	detachment	16	The lower part of the four walls of the Main chamber deterioration, detachme	No	2024
16	Main chamber	West wall of main chamber	detachment	7.5	Lower deterioration and detachment7.5 m 2	No	2024
16	Main chamber	North wall of main chamber	detachment	4.8	Whole wall sporadic flaking	No	2024

Figure 11. The manual verification results take Cave 15 deterioration increase and Cave 16 deterioration repair as examples.

Leveraging AI and graph databases enables multi-dimensional analysis of Mogao Grottoes wall paintings conservation data. Spatially, quantitative data is derived from cave-deterioration location node relationships; temporally, data across periods is analyzed to understand deterioration trends, predict development, and adjust conservation strategies promptly for long-term mural protection. Through case studies and data analysis, the feasibility and effectiveness of integrating graph-structured data with AI models in Mogao wall paintings conservation data management are validated.

4. Conclusions

This study addresses data management challenges in the conservation survey of Mogao Grottoes wall paintings by proposing an innovative approach integrating graph data structures with large AI models, effectively overcoming bottlenecks in traditional data management and analysis to establish a new paradigm for digital cultural heritage conservation.

Technically, AI tools play a crucial role in extracting and structuring multi-source data, while graph data structures model caves, walls, and deteriorations as nodes with spatio-temporal edges, enhancing data management efficiency. The synergy of graph data and AI enables multidimensional analysis of wall paintings deterioration information. In terms of application value, this approach improves the accuracy of conservation decision-making, allowing conservators to access deterioration information promptly and prioritize interventions in areas with rapid progression. It also sets a benchmark for data-driven management in cultural heritage, fostering digital transformation in the field.

However, we have also found that the non-uniformity of data record formats and the lack of detailed data have a significant impact on the analysis. In future verification work, we must standardize the requirements for records. Moreover, this approach still has room for improvement, especially in enhancing the retrieval efficiency of complex queries. Future research will focus on optimizing natural language processing models, integrating technologies such as image recognition and sensor monitoring, and promoting interdisciplinary collaboration to further advance the conservation and long-term preservation of cultural heritage.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-M-2-2025-327-2025

Acknowledgements

Corresponding author: Yipu Gong, email: gongyp@dha.ac.cn. This research was supported by the National Key R&D Program of China (Grant No. 2024YFF0907700) and the project at the institutional level of Dunhuang Academy (No.2024-KJ-YB-1).

References

Allegra, D.F., Andrea, Borghesi, Andrea, Boscarino, Michela, M., 2022. HADA: An automated tool for hardware dimensioning of AI applications. Knowledge-Based Systems Volume 251. https://doi.org/10.1016/j.knosys.2022.109199.

Asmaa, S., Mahmoud, H., Michael, K., Sarah, E., Youtham, B., Nermin, N., n.d. An AI Based Automatic Translator for Ancient Hieroglyphic Language—From Scanned Images to English Text.

Chen, G., 2017. Study on Salting Damage Analysis and Treatment of wall paintings at Mogao Grottoes, Dunhuang (PhD dissertation). Lanzhou University.

Cheng, F., Zhang, Hong, Fan, Wenjie, Harris, Barry, 2018. Image Recognition Technology Based on Deep Learning. Wireless Personal Communications 102, 1917–1933. https://doi.org/10.1007/s11277-018-5246-z

Dubois, C., Eigen, D., Delmas, E., Einfalt, M., Lemaçon, C., Berteloot, L., Bossuyt, P.M., Drummond, D., Scherdel, P., Simon, F., Torchin, H., Vali, Y., Bloch, I., Cohen, J.F., 2024. Deep learning in medical image analysis: introduction to underlying principles and reviewer guide using diagnostic case studies in paediatrics. BMJ e076703. https://doi.org/10.1136/bmj-2023-076703

Fan, J., 2000. Protection and Management of Mogao Grottoes in Dunhuang. Dunhuang Research 1–4. https://doi.org/10.13584/j.cnki.issn1000-4106.2000.01.001.

Gong, Y., Wang, X., Zhang, Z., Wang, S., Jin, K., n.d. A Large-Scale Time Series Data Management Platform for Cultural Heritage Site Risk Monitoring: Architecture Design and Application. ACM Journal on Computing and Cultural Heritage. https://doi.org/10.1145/3718957

Li, H., Wang, W., Zhan, H., Qiu, F., An, L., 2010. New judgement on the source of soil water in extremely dry zone. Acta Ecologica Sinica 30, 1–7. https://doi.org/10.1016/j.chnaes.2009.12.001

Li, J., Zhang, Hui, Fan, Z., He, X., He, S., Sun, M., Ma, Y., Fang, S., Zhang, Huabing, Zhang, B., 2013. Investigation of the renewed deteriorations on wall paintings at Mogao Grottoes. heritage science 1, 31. https://doi.org/10.1186/2050-7445-1-31

Menotti, G., 2025. The model is the museum: generative AI and the expropriation of cultural heritage. AI & Soc s00146-025-02290–1. https://doi.org/10.1007/s00146-025-02290-1

Pei, S., He, Y., 2020. Study on the Importance of Protecting the Original Artistic Value of the Cultural Heritage of the Mogao Grottoes, in: Proceedings of the 4th International Conference on Art Studies: Science, Experience, Education (ICASSEE 2020). Presented at the 4th International Conference on Art Studies: Science, Experience, Education (ICASSEE 2020), Atlantis Press, Moscow, Russia. https://doi.org/10.2991/assehr.k.200907.047

Wang, L., Ruan, W.-Q., Li, Y.-Q., 2025. Is AI heritage tourism interpretations better at deepening your cultural memory? Journal of Hospitality and Tourism Management 63, 68–76. https://doi.org/10.1016/j.jhtm.2025.03.011

Yi, J., Tian, Y., Zhao, Y., 2024. Novel Approach to Protect Red Revolutionary Heritage Based on Artificial Intelligence Algorithm and Image-Processing Technology. Buildings 14, 3011. https://doi.org/10.3390/buildings14093011