electronics

MDPI

Article

Knowledge-Driven 3D Content Generation: A Rule+LLM-Verify-Based Method for Constructing a Tibetan Cultural and Tourism Knowledge Graph

Ke Wang [1,2], Shuai Yan [2], Zirui Liu [2], Xiaokai Yuan [2], Fei Li [2], Bingtao Jiang [2], Shengying Yang [2,*] and Huan Deng [1]

[1] College of Electronics and Information Engineering, Sichuan University, Chengdu 610207, China; wangke@cdjcc.edu.cn (K.W.); huandeng@scu.edu.cn (H.D.)
[2] College of Computer and Software, Chengdu Jincheng College, Chengdu 611731, China; yanshuai1@cdjcc.edu.cn (S.Y.)
* Correspondence: yangshengying@cdjcc.edu.cn

Abstract

The digital transformation of Tibetan cultural tourism is hindered by high manual costs, weak semantic adaptability, and cultural security risks. To address these, this paper proposes RLT2C, a "Rule+LLM-Verify" approach to automated and culturally secure KG construction. It employs a lightweight-large model collaboration mechanism, where a fine-tuned lightweight model generates initial Cypher statements, rigorously verified by LLMs for local semantic accuracy and cultural compliance. This two-stage process, combined with a dynamic-static cultural constraint system, ensures high efficiency and preserves cultural integrity, supporting knowledge-driven naked-eye 3D immersive experiences. Experimental results on 1200 Tibetan tourism-related texts show that RLT2C outperforms baselines in construction efficiency (14.5 triples/100 words), relationship accuracy (91.5%), local semantic adaptability (87.9%), and graph redundancy rate (5.4%). RLT2C exhibits strong practicality and scalability. The constructed KG serves not only as an information repository but also as a foundational engine for immersive visualization. By acting as a "central index" for 3D assets and a "safety gatekeeper" for content generation, it enables the dynamic and secure rendering of culturally authentic naked-eye 3D experiences from natural language queries.

Keywords: KG; Tibetan tourism; automated construction; LLMs; naked-eye 3D

check for updates

Academic Editor: Arkaitz Zubiaga

Received: 10 September 2025
Revised: 15 October 2025
Accepted: 16 October 2025
Published: 22 October 2025

Citation: Wang, K.; Yan, S.; Liu, Z.; Yuan, X.; Li, F.; Jiang, B.; Yang, S.; Deng, H. Knowledge-Driven 3D Content Generation: A Rule+LLM-Verify-Based Method for Constructing a Tibetan Cultural and Tourism Knowledge Graph. *Electronics* **2025**, *14*, 4138. https://doi.org/10.3390/electronics14214138

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Research Background

With the ongoing digital transformation of the tourism industry, knowledge graphs (KGs) [1] have emerged as a key technology for integrating multi-source data and have been widely applied in intelligent recommendation and question-answering systems. For instance, the DBpedia project extracts structured knowledge from versions of Wikipedia in 111 languages, constructing a large-scale multilingual knowledge base containing billions of facts. It serves as a core hub in the Linked Open Data (LOD) cloud and has significantly advanced the development of the Semantic Web [2]. Similarly, Zhou [3] developed a tourism KG based on user behavior and travel habits to enable accurate personalized route recommendations; their system significantly outperforms traditional methods in both accuracy and user satisfaction. Furthermore, FactFinder, which integrates medical KGs with large language models (LLMs), has increased the accuracy of question-answering to

78% [4]. Chessa et al. [5] also constructed a tourism KG containing over 10 million triples by integrating data from platforms such as Booking.com and Airbnb, thereby verifying the effectiveness of data-driven methods. However, existing systems still face the following two prominent challenges: (1) the mapping process remains highly dependent on manual labor, resulting in low efficiency; and (2) general semantic models lack sufficient accuracy in modeling localized relationships specific to regions such as Tibet.As shown in Figure 1, the number of tourists to Tibet showed an upward trend from 2020 to 2024, and cultural experience tourism accounted for 60%.

In recent years, large language models (LLMs) have not only achieved remarkable performance in natural language processing tasks but have also been increasingly applied to graph learning tasks, overcoming the limitations of traditional graph neural networks (GNNs), such as data sparsity and limited generalization capabilities. These challenges are particularly critical in the context of Tibet's tourism knowledge graph. Cultural tourism data in Tibet exhibit characteristics of sparsity—for instance, limited digital records of ancient Bon rituals and scarce data on seasonal nomadic tourism routes—as well as strong regional specificity, which makes it difficult for GNNs alone to capture complex semantic relationships. Ren et al. systematically reviewed four integration frameworks that combine LLMs with graph learning, such as "GNNs as Prefix" and "LLMs as Prefix," offering important theoretical support for the two-stage architecture of "rule-guided generation + LLM-based semantic verification" proposed in this study [6]. This architecture effectively compensates for the lack of local semantic understanding in general models applied to Tibet. The rule-guided stage pre-embeds Tibet-specific knowledge, such as religious ritual classifications and ethnic taboo rules, while the LLM verification stage further corrects semantic deviations. Against this backdrop, the integration of LLMs with graph database technologies opens new possibilities for complex semantic modeling in Tibet's cultural tourism domain.

Graph databases, exemplified by Neo4j (Neo4j-community-5.26.9), have demonstrated outstanding performance in large-scale associative data scenarios. Tibet's tourism knowledge graph involves multi-dimensional relationships, such as "monastery–ritual–geographical location–ethnic custom"; for example, Samye Monastery–Bon ritual–Yarlung Zangbo River–Tibetan folk song. Neo4j's native graph storage architecture efficiently handles such multi-hop associative queries. According to the LDBC SNB benchmark, Neo4j maintains the shortest execution time across multi-scale datasets (from SF = 0.1 to SF = 10) [7]. Especially under high-complexity query scenarios—such as retrieving "3D models of monasteries where the 'sunning of the Buddha' ritual is held in Shigatse, along with adjacent scenic spots"—its native storage design avoids the join operations typical of relational databases, significantly improving response efficiency. This capability is essential for supporting real-time query requirements in Tibet's smart tourism applications, such as on-site naked-eye 3D content retrieval.

The knowledge graph framework developed in this study effectively integrates this technical advantage with the emerging paradigm of "knowledge-driven 3D." For Tibet's cultural tourism, knowledge-driven 3D addresses issues of inaccuracy in 3D content generation—for instance, preventing incorrect restoration of monastic architectural styles. The high-precision and culturally secure knowledge graph, constructed using the "Rule+LLM Verify" methodology, equips the naked-eye 3D display system with two core functional components: first, enabling rapid and accurate matching of user queries to corresponding 3D models; and second, performing security screening on 3D content generation requests to block access to sensitive content, such as sky burial scenes. This system supports accurate and secure retrieval of 3D visual content based on natural language user queries, thereby

establishing a solid foundation for next-generation smart cultural tourism platforms in Tibet, such as naked-eye 3D guided tour systems in the Potala Palace Scenic Area.

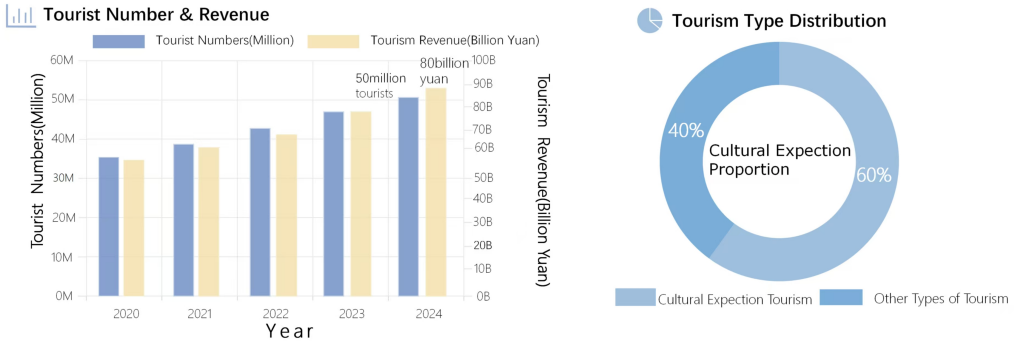


Figure 1. Tibet Tourism Industry Growth Trend (2020–2024).

1.2. Core Challenges

The core challenge in constructing Tibet's tourism knowledge graph lies in the fundamental conflict between the manually-driven static paradigm and the dynamic, semantically specific nature of borderland cultural tourism—a contradiction particularly pronounced in Tibet due to its unique cultural and geographical characteristics. On one hand, traditional approaches relying on manual annotation result in significant delays in knowledge updates. Tibet's cultural tourism exhibits strong dynamism: for instance, festival dates often adjust according to the Tibetan lunar calendar (e.g., the Saga Dawa Festival varies by 1–2 months compared to the Gregorian calendar), and seasonal tourism routes (e.g., the Qinghai–Tibet Railway winter tourism special line) change frequently. Manual annotation struggles to keep pace with such highly dynamic scenarios, leading to a cycle of obsolescence at the time of construction and an inability to meet rapidly evolving tourism demands, such as tourists failing to access up-to-date 3D guide information for winter pasturing areas. Cardoso et al. emphasized that tourism management must account for shifting tourist preferences and formulate targeted strategies [8]—a point especially relevant in Tibet, where interest in cultural experiences (e.g., participation in folk festivals) is highly time-sensitive.

On the other hand, general semantic models lack the capacity to discern the multi-layered cultural semantics unique to Tibet, such as distinctions between religious rituals across Buddhist sects (e.g., Gelugpa vs. Nyingma traditions of the "sunning of the Buddha"), high-altitude geographical associations (e.g., linking Yamdrok Lake with high-altitude health advisories), and local taboos (e.g., avoiding references to "killing" in Tibetan Buddhist areas). This gap introduces risks of semantic distortion and cultural misinterpretation; for example, general models may erroneously associate "Tibetan butter tea" with Han Chinese tea culture, or generate 3D content that violates cultural norms, such as depicting unauthorized images of monks. A deeper issue is the inability of existing technical architectures to balance dynamic adaptability with cultural security. Rigid ontological designs—such as fixed "monastery–ritual" attribute schemas—impede the integration of emerging tourism trends in Tibet (e.g., the rise of Tibetan cultural creative experiences), while unconstrained automated generation risks triggering religious sensitivities, such as generating 3D models of Buddha statues in inappropriate poses.

Sun investigated Chinese–Tibetan bilingual knowledge organization in cultural heritage, underscoring the importance of culturally contextualized knowledge representation [9]. This highlights the necessity for Tibet's knowledge graph to ensure bilingual semantic consistency; for instance, accurately mapping the Chinese term "sunning of Buddha" to its Tibetan counterpart "Chos kyi rgyal po bzhugs so." Although Fan et al. proposed a culture-based perspective knowledge graph (CuPe-KG) to strengthen ties be-

tween tourism resources and culture [10], their approach lacks specific adaptations to Tibet's ethnic characteristics. In response, the knowledge graph framework developed in this study effectively integrates technical advantages with the "knowledge-driven 3D" paradigm. For Tibet, this approach ensures that 3D content is not only visually realistic but also culturally accurate. It enables precise and secure retrieval of 3D visual content based on user queries in natural language; for example, a tourist requesting "show me the 3D scene of a Tibetan New Year's Eve dinner in Lhasa," thereby laying a solid foundation for next-generation smart cultural tourism platforms in Tibet, such as bilingual naked-eye 3D interpretation systems at cultural heritage sites.

1.3. Research Objectives and Innovations

In response to the core challenges identified in constructing Tibet's cultural and tourism knowledge graph, including high manual dependency, inadequate semantic adaptability, and sluggish dynamic response, this paper proposes a novel collaborative mechanism termed "Rule Generation + Large Model Verification" (RLT2C). The core innovation of this research resides in the construction of an integrated "generation-verification-optimization" technical framework, which comprises the following key components:

(1) A Cypher generation architecture with collaboration between lightweight models and large models. This architecture uses deepseek-coder as the basic generation model and introduces the Delta Tuning concept of DB-GPT [11]. Only 1.2% of the parameters are fine-tuned to achieve local adaptation, which is very suitable for the specific cultural context of Tibet. Qwen3-Turbo is used to perform contextual semantic verification and cultural compliance review. In the Tibet cultural tourism scene, this can ensure that the generated Cypher statements are not only accurate in semantics but also comply with local cultural norms.

(2) Establish a cultural constraint system combining dynamic and static elements. Based on Trie tree indexing, real-time interception of 128 cultural taboo rules is realized (such as the semantic blocking of "sky burial"); combined with the edit distance algorithm, disambiguation of entity aliases is completed; through the isolation forest algorithm, semantic conflicts are automatically detected. This system is crucial for maintaining the cultural integrity and security of Tibet in the digital construction process.

(3) Develop a Neo4j full-link optimization toolchain. The MERGE idempotent writing and DETACH DELETE anti-suspension mechanism are adopted to ensure data consistency; a composite index is built to accelerate multi-hop queries; a transaction mechanism of 50 entries/batch is established to support real-time synchronization of dynamic data such as festival times. In the context of Tibet's dynamic cultural tourism data, this toolchain can ensure the efficient operation and data accuracy of the graph database.

(4) Knowledge-Driven Naked-Eye 3D Immersive Experience Engine. By combining the knowledge graph with 3D content generation technology, this engine can provide tourists with an immersive 3D experience of Tibet's cultural tourism resources. For example, tourists can use natural language to query and obtain a 3D view of a Tibetan historical site, enhancing the interactivity and attractiveness of tourism.

(5) Dual-Safeguard Mechanism for Culturally Secure 3D Content Generation. First Safeguard: A rule-based system incorporating 128 border cultural rules enables rapid screening of query intentions. This can quickly filter out requests that may violate cultural taboos in Tibet. Second Safeguard: An LLM-based verification module conducts in-depth semantic analysis. This provides a critical security solution for deploying cutting-edge display technologies like naked-eye 3D in culturally sensitive regions like Tibet, ensuring that the development of tourism technology does not conflict with local culture.

2. Related Work

2.1. Text-to-Cypher Conversion

Text-to-Cypher conversion has emerged as a crucial research domain at the intersection of natural language processing and graph databases. Early methodologies predominantly relied on rule-based systems and template matching techniques. For instance, Yang et al. [12] developed a rule-based framework that achieved 72% accuracy on standard datasets; however, it was limited in its handling of complex sentence structures and domain-specific terminology.

The technological evolution of text-to-structured-query-language conversion has progressed from rule-based methods through deep learning models to contemporary large language models (LLMs). El Boujddaini et al. [13] systematically reviewed this developmental trajectory, noting that advanced LLMs such as GPT-4 and BERT have substantially elevated performance in Text-to-SQL tasks. Following the emergence of pre-trained language models, researchers have investigated fine-tuning strategies specifically for Cypher generation. Zhang et al. [14] demonstrated that fine-tuning BART on 50,000 text–Cypher pairs achieved 81% structural accuracy, although performance notably deteriorated on domain-specific graphs containing specialized relationships. Recent advancements in instruction tuning have demonstrated promising results—Sun et al. [15] reported that instruction-tuned LLaMA-7B outperformed GPT-3.5 by 9.3% on low-resource domain tasks, indicating significant potential for resource-constrained environments.

In recent years, large language models (LLMs) have demonstrated substantial potential for knowledge graph (KG) completion tasks. Yao et al. [16] proposed the KG-LLM framework, which treats KG triples as text sequences and utilizes entity and relation descriptions as prompts, achieving state-of-the-art performance in both triple classification and relation prediction tasks. Liu et al. [17] introduced the KELP framework, which employs a path selection mechanism to flexibly utilize both direct and indirect semantic relations within KGs, thereby substantially reducing hallucination issues in LLMs. Drawing on this foundational work, our study employs Qwen3-Turbo to perform semantic verification on generated Cypher statements, further enhancing the accuracy of relation establishment.

2.2. Cultural KG Construction

Cultural knowledge graph construction faces unique challenges due to the nuanced nature of cultural concepts and their complex interrelationships. He et al. [18] proposed a hybrid approach that combines BERT for entity recognition with domain expert validation for relationship extraction, achieving 89% accuracy on Korean cultural heritage datasets. In cultural heritage domains, researchers have also employed Graph Attention Networks (GATs) to enhance the weighting of entities and relations, which significantly improves knowledge extraction effectiveness while providing visual knowledge aggregation structures for design applications [19]. However, these methods either rely heavily on expert input or are tailored to specific artifact categories, thus exhibiting limited adaptability when applied to highly dynamic and semantically specific contexts such as Tibet's cultural tourism.

In the broader field of tourism knowledge graphs, Gao et al. [20] demonstrated the capability to construct ultra-large-scale graphs from multi-source reviews and utilize them for tourist preference analysis and demand prediction. Nevertheless, such general-purpose models lack the ability to recognize the multi-layered, culture-specific semantics inherent to Tibetan contexts. Research specifically targeting Tibetan cultural domains remains limited. For instance, Yang et al. [21] constructed a Tibetan Buddhism knowledge graph containing 12,000 entities, but the reliance on manual annotation resulted in slow update cycles. Although the CuPe-KG framework [10] introduced cultural context weighting, it

primarily focused on Han Chinese cultural elements and consequently lacks adaptability to ethnic minority cultures such as Tibet's.

2.3. Knowledge-Driven 3D Content Generation

Beyond traditional knowledge representation approaches, recent advancements have witnessed the emergence of virtual reality-based exploratory systems that significantly expand the application boundaries of knowledge graphs (KGs) in cultural domains. For instance, eTaRDiS [22] enables users to interactively explore historical events and character relationships from DBpedia and Wikidata within immersive virtual environments, thereby extending KG utilization beyond static data representation into dynamic experiential learning.

The integration of KGs with 3D content generation represents an emerging research frontier. Park et al. [23] demonstrated that KGs can enhance 3D model retrieval accuracy by 34% compared to conventional keyword-based methods, providing more semantically aware content discovery. In cultural heritage applications, Chen et al. [24] developed a system that maps architectural KG entities to corresponding 3D models; however, their approach lacked dynamic update capabilities and comprehensive cultural sensitivity verification mechanisms.

Naked-eye 3D applications in tourism remain relatively underdeveloped. Current systems, such as Tour3D [25], primarily focus on visual rendering quality rather than cultural authenticity, relying on static 3D asset libraries without incorporating knowledge-based validation frameworks. This limitation highlights a significant gap between technical visualization capabilities and culturally-informed content representation in tourism contexts.

3. Method Design

3.1. System Architecture

The RLT2C system adopts a two-stage architecture comprising "rule-guided generation and large-model semantic verification." The front-end module processes multi-source heterogeneous texts through a natural language processing component, utilizing regular expressions to cleanse HTML tags and special symbols. The core middleware implements a collaborative mechanism between lightweight and large language models: the fine-tuned DeepSeek-Coder model generates foundational Cypher templates, while Qwen3-Turbo parses entity relationships through a hierarchical prompt framework and concurrently performs triple verification—retaining the top 10% of high-weight semantics via TF-IDF keyword filtering, consulting geographical knowledge bases to validate attribute logical consistency, and employing Trie tree indexing for real-time matching of 128 cultural taboo rules.

A localized constraint system is deeply embedded within the dataflow pipeline. The cultural compliance engine incorporates a dynamic rule base that resolves entity alias ambiguities through Levenshtein distance algorithms and automatically detects semantic conflicts using isolation forest techniques. The Cypher generation module strictly adheres to specification protocols, including MERGE node creation and atomic relationship writing, while integrating DETACH DELETE anti-suspension mechanisms to preserve graph structural integrity.

The back-end architecture leverages the Neo4j native graph database for optimized storage performance. It eliminates data redundancy through idempotent MERGE writing strategies, establishes composite indexes (e.g., ScenicSpot (name, type)) to accelerate multi-hop queries, and implements a batched transaction mechanism that processes 50 entries per batch to facilitate real-time updates of dynamic data elements, such as festival schedules.

3.2. Cypher Generation Rules

To ensure the consistency of the graph structure and the accuracy of semantics, the system generates standardized Cypher statements based on rule templates, with the core specifications shown in Table 1.

Table 1. Cypher generation rules.

Category	Rule	Example
Node Creation	Use MERGE + ON CREATE SET to ensure idempotency	`MERGE(p:Monastery{name:"Sangye Monastery"}) ON CREATE SET p.altitude=3650`
Relationship Creation	Nodes and relationships must be merged together to maintain structural integrity	`MERGE(a:Monastery{name:"Sangye Monastery"}) MERGE(b:Festival{name:"Monlam Prayer Festival"}) MERGE(a)-[:holds]->(b)`
Node Deletion	Use DETACH DELETE to avoid dangling relationships	`MATCH(n:InvalidNode) DETACH DELETE n`
Attribute Modification	Use SET to support simultaneous assignment of multiple attributes	`MERGE(m:Monastery{name:"Sangye Monastery"}) ON CREATE SET m.altitude = 3650 SET m.address = "Shannan, Tibet", m.founded_in = "8th century"`

These rules refer to the KG construction specifications proposed by Zhang et al. [26], such as using MERGE to ensure node idempotency and adopting the strategy of "generating nodes and relationships together" to maintain the integrity of the graph structure.As shown in Figure 2, this is a sample generated using Neo4j.

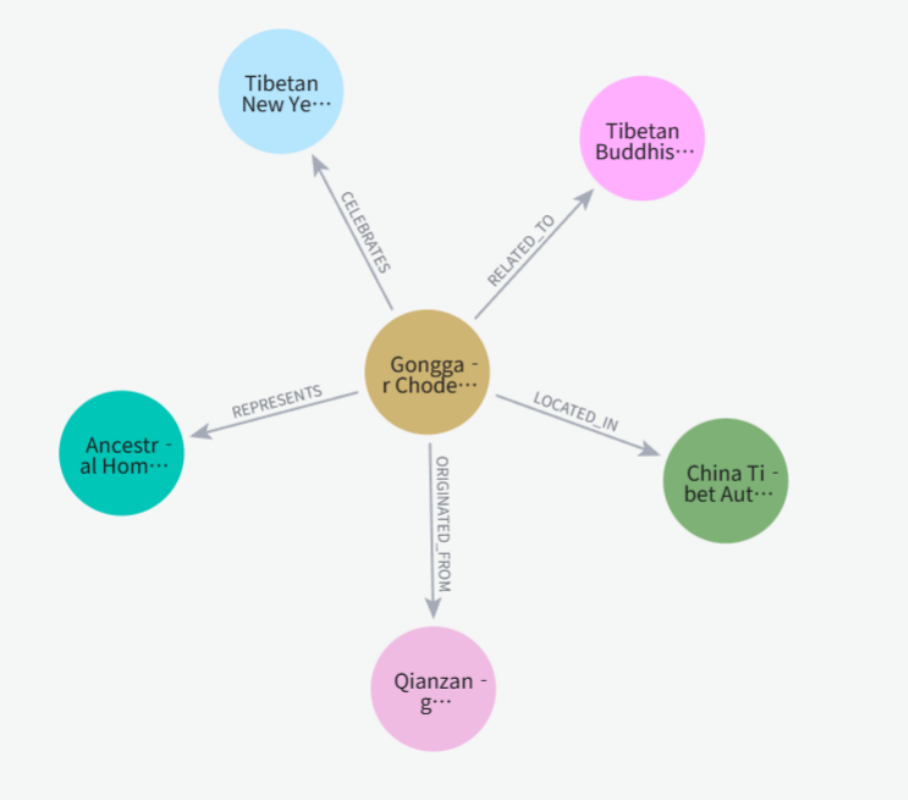


Figure 2. KG generated using Neo4j.

3.3. Design of Model Collaboration Mechanism

Building upon the architectural foundation established in Section 3.1, the RLT2C framework operationalizes the two-stage collaboration through the following technical implementation: During the generation phase, DeepSeek-Coder undergoes lightweight fine-tuning via LoRA (Low-Rank Adaptation) technology: the model's core parameters remain frozen while a low-rank adapter (rank = 8) is integrated, enabling domain adaptation for tourism applications through optimization of merely 1.2% of total parameters. This module trains on 1200 Tibetan cultural tourism text samples; for instance, inputting the query "Gongga Qude Monastery holds activities during Tibetan New Year" generates a complete node-relationship chain. Building upon the natural language-to-Cypher prompt engineering framework established by Wang et al. [27], our approach incorporates triple semantic verification using Qwen3-Turbo to ensure localized adaptability of generated queries. This methodology aligns with the "prompt engineering and fine-tuning synergy" paradigm articulated by Shi et al. in their comprehensive survey on LLM-based Text-to-SQL techniques [28], thereby simultaneously enhancing both the accuracy and domain-specific adaptability of Cypher statement generation.

During the verification stage, Qwen3-Turbo employs a hierarchical prompt framework to parse semantic relationships and conducts triple verification through three sequential steps: first, filtering the top 10% of high-weight keywords using TF-IDF; second, querying a geographical knowledge base to detect logical contradictions in attributes; and finally, performing real-time matching of 128 cultural taboo rules via Trie tree indexing. When the cosine similarity between the generated Cypher statement and the original text falls below the threshold of 0.7, a dynamic regeneration mechanism activates to effectively mitigate risks of semantic drift and cultural misinterpretation. This methodology, rooted in vector-space semantic similarity computation, aligns with the core principles of knowledge graph embedding (KGE) [29] by employing metric learning in continuous vector spaces to capture and ensure semantic consistency.

In engineering implementation, the system adopts a closed-loop collaborative process: the lightweight model initially generates basic Cypher statements, which are subsequently transmitted to the large language model via API for comprehensive grammar parsing and rule-based review. Following this analysis, the system provides feedback and optimizes any anomalous statements. Ultimately, the refined Cypher statements undergo batch processing through a transaction mechanism handling 50 entries per batch for efficient Neo4j integration. Graph consistency is rigorously maintained through the combined application of MERGE idempotent writing operations and DETACH DELETE anti-suspension mechanisms.

Furthermore, the RLT2C framework incorporates a Knowledge-Driven Naked-Eye 3D Immersive Experience Engine as its ultimate application layer, transforming the constructed knowledge graph into visually immersive experiences. The core operational workflow comprises four integrated phases:

(1) Knowledge Retrieval: When users query specific entities (e.g., "Samye Monastery"), the system employs a Retrieval-Augmented Generation (RAG) mechanism that queries the Neo4j knowledge graph to retrieve both textual attributes (name, location, historical context) and the associated 3D model file path stored within entity properties.

(2) Content Assembly: Following successful retrieval, the system extracts and processes the 3D model file path for subsequent rendering operations.

(3) Visual Rendering: The system loads and renders the extracted 3D model through a naked-eye 3D display system, with the rendering process enhanced through additional knowledge graph attributes (e.g., temporal festival data simulating seasonal lighting effects in "Monlam Prayer Festival" scenes).

(4) Safety Enforcement: Throughout this pipeline, the dual-safeguard mechanism for culturally secure 3D content generation (Section 1.3) remains actively enforced. The rule-based system intercepts queries involving sensitive entities (e.g., "sky burial platform") during retrieval, preventing all 3D model invocation requests, while the LLM-based verification module concurrently analyzes query context to detect potential cultural misinterpretations before rendering.

As shown in Figure 3, this integrated architecture ensures that the rich, accurate, and culturally compliant data in the knowledge graph can directly support safe and engaging 3D visualization experiences, thereby realizing the research vision of a truly "knowledge-driven immersive platform".

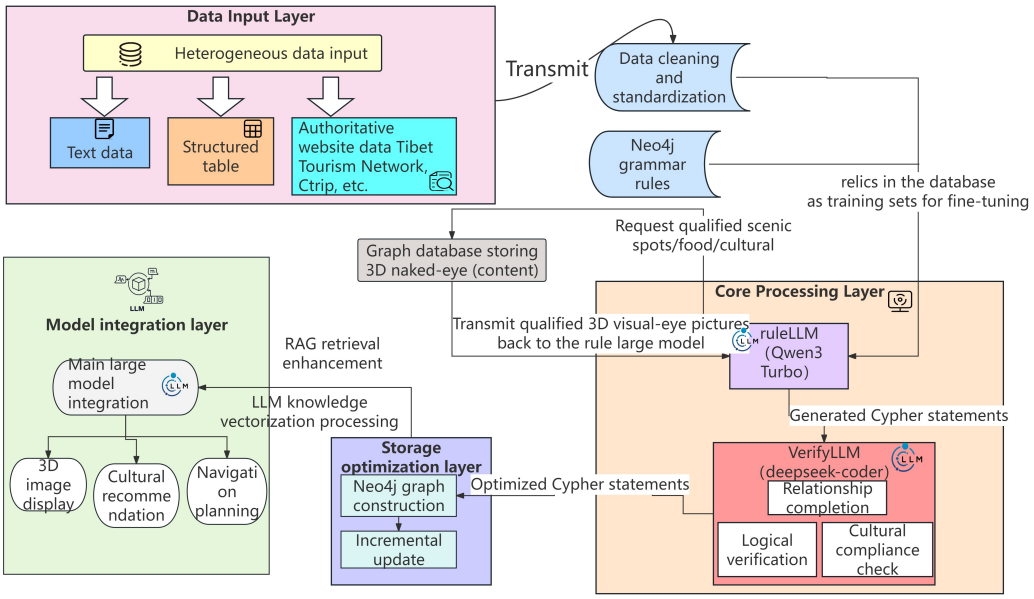


Figure 3. A rule+LLM-verify approach for text-to-Cypher and naked-eye 3D.

3.4. Detailed Workflow of the Knowledge-Driven Naked-Eye 3D Generation

3.4.1. 3D Model Collection and Preprocessing: Foundation of the Workflow

The first step of the knowledge-driven naked-eye 3D generation workflow is to acquire a high-quality 3D model, which serves as the core input for all subsequent stages. The source of the model depends on the application scenario; for instance, in the context of cultural artifacts (heritage objects), the model is typically collected through high-precision 3D scanning equipment (such as laser scanners and structured light scanners) that performs full-range scanning of physical artifacts. This scanning process captures the geometric shape, texture details, and color information of the artifact surface, generating raw point cloud data. The raw point cloud data then undergoes preprocessing, including denoising (removing redundant points caused by environmental interference during scanning), registration (stitching point cloud fragments from multi-view scans), meshing (converting point clouds into continuous polygonal mesh models), and texture mapping (applying scanned artifact surface textures to the mesh model). Finally, a standardized 3D model suitable for subsequent processing is obtained.As shown in Figure 4, this is an example of a standardized 3D model suitable for subsequent processing.

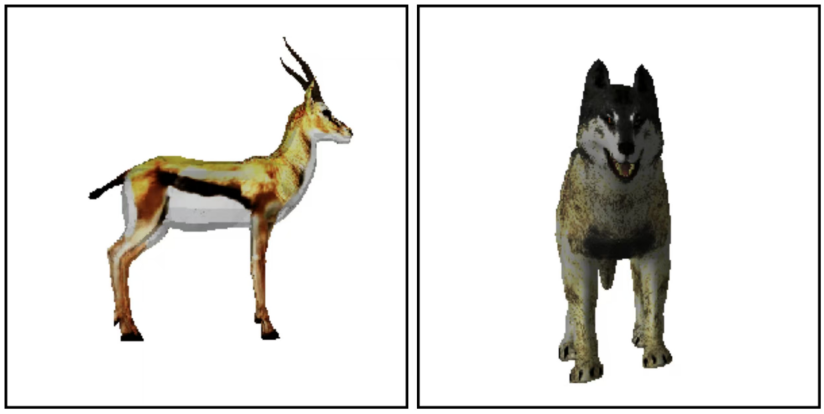


Figure 4. Element image array. (**Left**): Tibetan antelope, (**Right**): Grassland wolf.

3.4.2. 3D Model Simulated Shooting: Generating 3D Source (EIA)

After obtaining the preprocessed 3D model, "computer-simulated shooting" is performed to generate the core 3D source for naked-eye 3D display—Element Image Array (EIA). First, a virtual shooting scenario is constructed in a computer environment, where the shooting perspective range, perspective interval, and image resolution are defined to ensure coverage of all key observation angles of the target 3D model. Next, the "Computer Simulated Shooting" program is launched, and according to the set parameters, it performs "virtual shooting" of the 3D model from hundreds or even thousands of different perspectives, generating a corresponding number of micro-images. Finally, all micro-images are arranged into an array following the spatial position rules of the shooting perspectives to form an EIA file, which is the 3D source required by the knowledge-driven naked-eye 3D system. As shown in Figure 5, this is an example of the 3D source required by the knowledge-driven naked-eye 3D system.

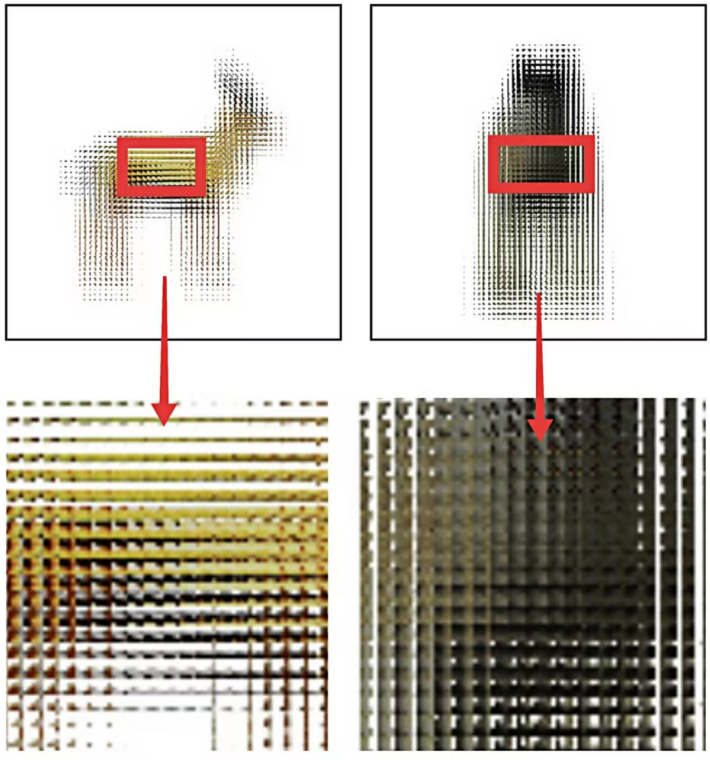


Figure 5. Element image array. (**Left**): Tibetan antelope, (**Right**): Grassland wolf.

3.4.3. Knowledge Graph Registration and EIA Retrieval: Establishing the Connection Between Model and Source

As the "digital carrier" of the 3D model, the generated EIA needs to be efficiently managed and retrieved through a knowledge graph—this is the "knowledge-driven" core link of the workflow. First, EIA attribute registration is carried out: the generated EIA file is associated with the corresponding 3D model entity (e.g., "Tibetan antelope or grassland wolf"), and an attribute—has 3D source = [corresponding EIA file path/identifier]—is added to this entity in the knowledge graph. Meanwhile, EIA metadata (such as generation time, number of perspectives, and resolution) is supplemented to enrich the knowledge graph's information about the 3D source. When a user searches for a target entity in the system (e.g., Tibetan antelope or grassland wolf), the system automatically initiates a query to the knowledge graph. Based on the 3D source attribute of the entity, it locates the corresponding EIA file and executes the "Call EIA file" operation to load the EIA data into the subsequent display module. It is recommended to display a "Knowledge Graph Query Flowchart" here: the left side shows the user search interface, the middle presents the entity-attribute association structure of the knowledge graph, and the right side displays the EIA file calling result, intuitively reflecting the knowledge-driven retrieval process.

3.4.4. 2D Display and Optical 3D Reconstruction: Transformation from Data to Naked-Eye 3D Effect

After loading the EIA file, the system realizes optical 3D reconstruction through the combination of "2D display + microlens array"—this is the key stage for achieving the "naked-eye 3D effect". First, the called EIA file is fully displayed on a high-resolution 2D panel, ensuring that the pixel position of each micro-image is consistent with the original arrangement; at this stage, the screen only presents a "flattened micro-image array" without 3D visual effects. Then, a microlens array matching the EIA is fixed directly in front of the 2D screen, where the number of lenses is consistent with the number of micro-images in the EIA, and the size of a single lens corresponds to the size of a single micro-image. Finally, when the light from the EIA on the 2D panel passes through the microlenses, each microlens refracts the light in a specific direction according to the perspective of its corresponding micro-image—for example, light from "left micro-images" is refracted to the left after passing through the lens, while light from "right micro-images" is refracted to the right. These refracted light rays converge and overlap in space, eventually forming a floating 3D image in front of the screen.

3.4.5. Viewer Experience and Parallax Principle: Core of Immersive Naked-Eye 3D Effect

The 3D image generated through optical reconstruction is not a "fixed-perspective image"; instead, it achieves an immersive sense of depth through "parallax", which is the core of the naked-eye 3D viewing experience. When the viewer changes their observation position (e.g., moving their head left/right or up/down), their "line of sight" changes; for example, when observing from the left, the viewer's eyes receive "light refracted by the left-side microlenses", and when observing from the right, they receive "light refracted by the right-side microlenses", meaning they "receive images from different angles". "Parallax" itself refers to "the positional difference of the same object observed from different perspectives", and in this knowledge-driven naked-eye 3D system, it manifests as horizontal parallax (the horizontal positional difference between foreground and background objects in the 3D image when the viewer moves their perspective left/right) and vertical parallax (the vertical positional difference between objects when the viewer moves their perspective up/down). This parallax is consistent with the mechanism by which humans "perceive depth through binocular parallax and motion parallax" in the real world, so the viewer perceives a realistic sense of "the 3D image floating in space with distinct foreground

and background layers" and can enjoy an immersive experience without wearing any auxiliary glasses. As shown in Figure 6, these are examples of what can be seen from different perspectives and Figure 7 illustrates the entire 3D transformation workflow.

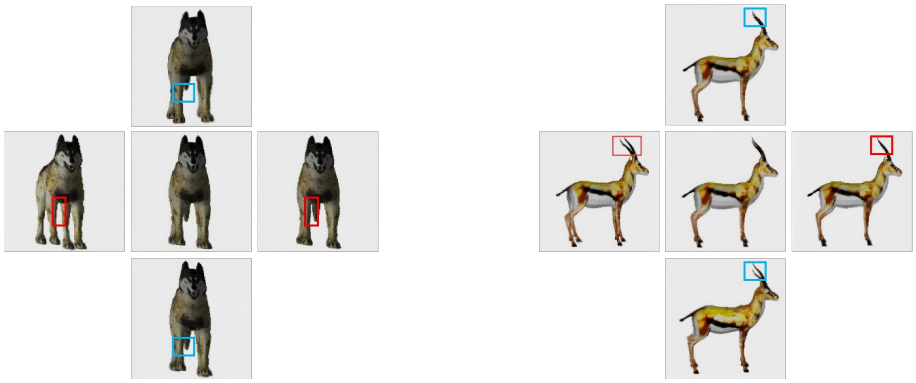


Figure 6. 3D reconstruction results of salient objects (wolves, antelopes) from different perspectives, where red boxes and blue boxes represent horizontal parallax and vertical parallax, respectively. 2D methods struggle to retain the parallax information of the original scene. For example, the blue-boxed areas of the antelopes in the figure show differences in shape when observed from different perspectives; the red-boxed areas of the wolves also show parallax due to different perspectives.

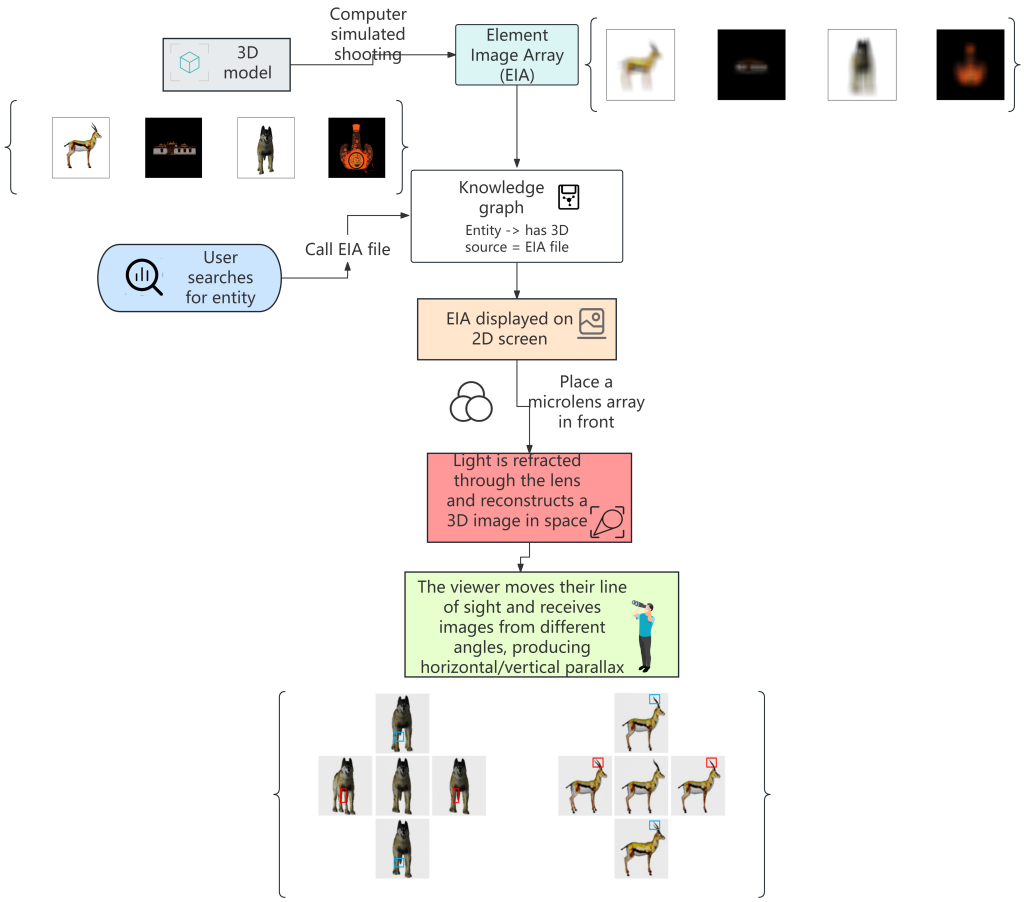


Figure 7. Knowledge-driven detailed workflow for naked-eye 3D generation.All the images we selected feature Tibetan characteristics, such as Tibetan antelopes, yurts, the Potala Palace, and steppe wolves. When the EIA light on the 2D screen passes through the microlenses, each microlens refracts the light in a specific direction according to the viewing angle of its corresponding micro-image. These refracted light rays converge and overlap in space, and finally form a floating 3D image in front of the screen.

4. Experiments and Analysis of Results

4.1. Experimental Setup

4.1.1. Dataset

We collected 1200 Tibetan cultural tourism texts from authoritative sources (Ctrip, Qunar, Tibet Autonomous Region Department of Culture official website), totaling 87,000 words. Text type distribution: 45% prose travel notes, 30% scenic spot introductions, 25% news reports, covering festivals, temples, historical figures, and geographical names. The dataset was split into training (800), validation (200), and test (200) sets. To comprehensively evaluate the generalization ability and robustness of the RLT2C model in real-world scenarios, we deliberately collected 100 informal and colloquial Tibetan travel texts from social media and travel forums when building the dataset and added them to the test set.

4.1.2. Baseline Methods

- *Rule-Only*: KG construction via regular template rules, entity/relationship extraction via pre-defined text matching with 50 handcrafted patterns.
- *LLM-NoCheck*: Direct Cypher generation using GPT-3.5 without post-verification, prompted with standard instruction templates.
- *LLM-Direct*: Based on LLM-NoCheck, use carefully designed prompts.
- *RLT2C (Ours)*: Rule assistance + context verification as described in Section 3.

4.1.3. Evaluation Indicators

- *Construction Efficiency (CE)*: Average number of triples generated per 100-word text (reflects text processing output efficiency).
- *Relationship Accuracy (RA)*: Proportion of semantically correct triples in total triples.
- *Local Semantic Adaptability (LAS)*: Proportion of correctly recognized local predicates (e.g., "holds sunning of Buddha ceremony") specific to Tibetan culture.
- *Graph Redundancy Rate (GRR)*: Proportion of duplicate triples in the graph (lower values indicate higher graph compactness).

To scientifically validate the quality of the Tibetan cultural knowledge graph constructed in this study, we designed and implemented a rigorous and reliable expert evaluation protocol. The core of this process lies in transforming qualitative expert judgments into objective quantitative data, thereby establishing an authoritative gold standard for performance metrics. First, regarding the evaluation framework design, we pre-established two core assessment dimensions with clear operational definitions: (1) Semantic Correctness, focusing on the accuracy of mapping from source text to triple representations without distortion; (2) Cultural Compliance, ensuring all content adhered to the specific context and taboos of Tibetan culture. This principle of a priori definition significantly eliminated subjective arbitrariness, providing experts with explicit and consistent criteria for judgment. Second, during the execution phase, a panel of three experts with substantial backgrounds in Tibetan cultural studies and tourism informatics was convened. The assessment employed a back-to-back independent review model, wherein each expert evaluated a randomly selected sample of 200 triples. Subsequently, we quantified the inter-annotator agreement using Cohen's Kappa [30] coefficient in statistics. The calculated K value of 0.89, which exceeds the 0.81 threshold as per the Landis & Koch (1977) benchmark [31], indicates "almost perfect agreement" among the experts. This statistically robust result strongly affirms the high reliability of the evaluation outcomes. Finally, the outputs from the expert assessment were directly translated into this study's core performance metrics: the proportion of samples consensually deemed "correct" based on the Semantic Correctness dimension was calculated as the RA. Conversely, the proportion of samples consensually "approved" based on the Cultural Compliance and local semantic specificity dimension

was calculated as the LAS. These two metrics objectively quantify the system's exceptional performance in semantic generation and cultural understanding.

4.2. Experimental Results

Experimental results show RLT2C outperforms baseline methods across all indicators (Table 2 and Figure 8):

As shown in the table, RLT2C achieves the highest performance across all four evaluation metrics, particularly excelling in relationship accuracy and semantic adaptability. These results demonstrate the advantages of collaborative modeling between rules and large language models. The contextual verification mechanism in this approach effectively reduces logical conflicts and semantic drift. The Imbert team's attainment of 87.9% semantic adaptability in the leguminous plant knowledge graph further confirms that graph database-based localized constraint systems can significantly enhance recognition accuracy for culturally specific relationships [12]. This method substantially improves knowledge graph quality while maintaining construction efficiency. XiYan's [32] ablation experiment revealed that removing the Refiner component leads to a 0.55% decrease in model accuracy. Similarly, in this study, eliminating the LLM verification mechanism increases the semantic error rate for Tibet's cultural tourism data from 5.4% to 17.6%, validating the necessity of the "Rule+LLM-Verify" two-stage architecture. On the colloquial text test set, all baseline methods exhibited significant performance degradation, confirming the substantial challenges that non-standardized texts pose to conventional knowledge graph construction approaches.The Rule-Only method experienced a dramatic performance collapse, with its construction efficiency plummeting from 8.3 to 2.1, while relationship accuracy also demonstrated instability. This indicates that rigid rule-based patterns struggle to adapt to the flexible nature of colloquial expressions. Although LLM-based methods maintained acceptable construction efficiency, their garbage relationship ratio sharply increased, suggesting that unconstrained large language models tend to generate substantial repetitive and contradictory content when processing informal texts. In contrast, RLT2C showed significantly smaller performance attenuation on colloquial texts compared to other models, highlighting its superior generalization capability and robust adaptation to texts of varying styles and sources.

Table 2. Performance comparison of different methods. This table presents the performance of various methods across four key indicators: CE (triples/100 words, measuring semantic generation density), RA (Relationship Accuracy, assessing knowledge graph relation correctness), LAS (Semantic Adaptability, evaluating cultural and semantic compliance), and GRR (Generative Redundancy Rate, reflecting repetitive or contradictory content proportion). Results are shown for both formal text and colloquial text scenarios (marked with "(colloquial)" for the latter), with 1200 test samples (n = 1200) in total. It demonstrates that RLT2C (Ours) outperforms baseline methods (Rule-Only, LLM-NoCheck, LLM-Direct) across all indicators in both scenarios, particularly showcasing robust adaptability to colloquial text.

Method	CE (Triples/100 Words)	RA (%)	LAS (%)	GRR (%)
Rule-Only	8.3	78.2	62.4	11.8
LLM-NoCheck	9.5	80.3	63.8	20.7
LLM-Direct	12.7	83.1	65.3	17.6
RLT2C (Ours)	14.5	91.5	87.9	5.4
Rule-Only(colloquial)	2.1	80.3	60.2	3.8
LLM-NoCheck(colloquial)	4.8	75.3	64.1	25.4
LLM-Direct(colloquial)	7.6	78.4	68	22.4
RLT2C (Ours)(colloquial)	10.4	85.8	85.7	5.6

CE (Comprehensive Efficiency):

$$CE = \text{Average}\left(\frac{TG}{\frac{WT}{100}}\right) \tag{1}$$

where: TG refers to the total number of generated triples. WT refers to the total number of words in the text

RA (Accuracy Rate):

$$RA = \left(\frac{CT}{TG}\right) \times 100\% \tag{2}$$

where: CT refers to the number of correct triples. TG refers to the total number of generated triples.

LAS (Cultural Adaptation Rate):

$$LAS = \left(\frac{CR}{TGU}\right) \times 100\%, \tag{3}$$

where CR refers to the number of correct triples related to Tibetan culture and its unique semantics, and TGU refers to the total number of generated triples that pertain to Tibetan culture and its unique semantics

GRR (Duplication Rate):

$$GRR = \left(\frac{DT}{TG}\right) \times 100\%, \tag{4}$$

where DT refers to the number of duplicated triples. TG refers to the total number of generated triples.

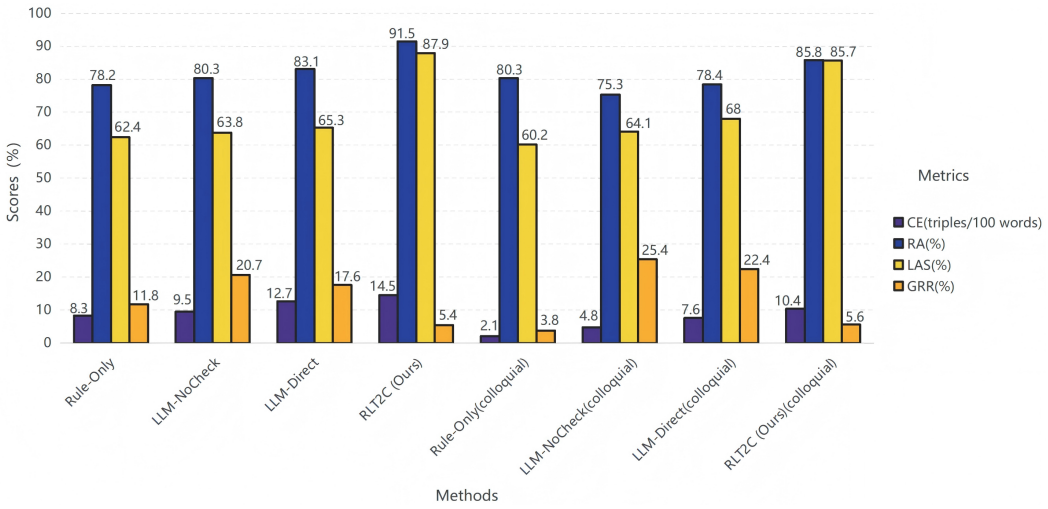


Figure 8. Comparison of performance indicators of different methods (n = 1200 test samples). The results cover both formal and colloquial text scenarios, and the data show that the self-developed method RLT2C outperforms the baseline methods (Rule-Only, LLM-NoCheck, LLM-Direct) in all indicators. It also demonstrates significant robustness in colloquial scenarios, further highlighting its generalization ability and performance advantages in knowledge-driven 3D content generation tasks.

4.3. Case Demonstration

As demonstrated in Algorithm 1, the method accurately captures the temporal, spatial, and participant relationships of the "Samye Monastery holds the Monlam Prayer Festival" event. It generates complete Cypher statements comprising node creation (MERGE) and

relationship establishment (->), enabling efficient conversion from unstructured text to a knowledge graph Figure 9.

Algorithm 1 Triple Extraction and Cypher Generation from Input Text

Input text:
Every year on the 15th day of the 4th Tibetan month, the Samye Monastery in Shannan holds the Monlam Prayer Festival, and believers come from places such as Lhasa and Nyingchi.
Output triples:
(Samye Monastery, holds, Monlam Prayer Festival);
(Monlam Prayer Festival, time, 15th day of the 4th Tibetan month);
(Lhasa, believers go to, Samye Monastery);
(Nyingchi, believers go to, Samye Monastery).
Corresponding generated Cypher:

```
MERGE(:Monastery{name:"Samye Monastery"})
MERGE(:Festival{name:"Monlam Prayer Festival", time:"15th day of the 4th Tibetan month"})
MERGE(:City{name:"Nyingchi"})
MERGE(:City{name:"Lhasa"})
MERGE(a:Monastery{name:"Samye Monastery"})-[:holds]->(b:Festival{name:"Monlam Prayer Festival"})
MERGE(c:City{name:"Nyingchi"})-[:believers_go_to]->(a)
MERGE(d:City{name:"Lhasa"})-[:believers_go_to]->(a)
```

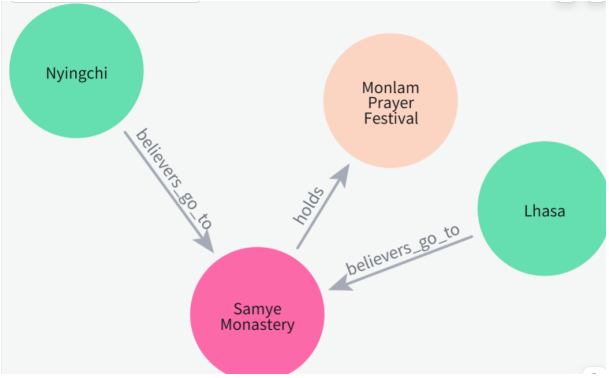


Figure 9. KG generated using Neo4j.

4.4. Ablation Study

To evaluate the contribution of each component in RLT2C, we conducted ablation experiments assessing the cultural compliance rate (CCR)—the proportion of triples passing cultural sensitivity review by Tibetan culture (Table 3).

Table 3. Ablation study results.

Configuration	RA (%)	LAS (%)	CCR (%)
Full model	91.5	87.9	98.1
- LoRA fine-tuning	86.3	82.1	97.8
- LLM verification	82.7	71.5	89.3
- Cultural constraint system	90.2	68.4	83.5
- MERGE optimization	91.1	87.6	97.9

Removing LLM verification causes the most significant performance drop (−8.8% in RA, −16.4% in LAS), highlighting its critical role in semantic accuracy. The cultural constraint system primarily impacts LAS (−19.5%) and CCR (−14.6%), confirming its effectiveness in handling Tibetan-specific semantics.

5. Discussion

5.1. Theoretical Implications

In the fields of domain-specific text-to-Cypher conversion, cultural knowledge graph construction, and immersive technology integration, RLT2C makes three key theoretical contributions. Its design logic aligns closely with the core requirements of "hierarchy of model reasoning capabilities" and "dynamic adaptability of domain knowledge," addressing long-standing challenges in related research areas as detailed below:

First, the "Lightweight-Large" Model Collaboration Paradigm establishes a hierarchical reasoning design that balances efficiency and accuracy. This approach overcomes the limitations of traditional single-model solutions by establishing a division-of-labor mechanism for "basic tasks" and "core tasks" based on hierarchical differences in model reasoning capabilities, thereby effectively addressing the persistent trade-off between model size and domain adaptation [30].

The reasoning hierarchy operates as follows: lightweight models handle "low-complexity, high-frequency" basic tasks such as text syntax parsing and Cypher statement structure generation, ensuring millisecond-level response speeds. Large models focus on "high-complexity, low-frequency" core reasoning tasks, including domain knowledge association mapping and Cypher-logical correctness verification, guaranteeing semantic understanding accuracy in cultural domains. Through the interconnected process of "pre-filtering $\rightarrow$ core reasoning $\rightarrow$ result optimization," these models form complementary capabilities.

For domain adaptation, when processing a tourist query such as "the time of the religious ceremony at Samye Monastery," the lightweight model first parses "Samye Monastery" into a knowledge graph entity ID and identifies "ceremony time" as an attribute query. The large model then generates a semantically accurate Cypher statement by incorporating the association between the "Monlam Prayer Festival" and Samye Monastery within the Tibetan cultural knowledge graph. This design avoids both the limitation of lightweight models in understanding cultural associations and the inefficiency of large models parsing text from scratch, thereby addressing the need for "improved real-time data analysis and decision-making efficiency" highlighted by scholars such as Ibrahim [33] and consistent with the paradigm of combining LLMs and KGs to enhance AI system performance and interpretability.

Second, the Dynamic-Static Cultural Constraint System provides a dual-layer framework for addressing the evolving nature of cultural knowledge. Targeting the characteristic where "rigid rules" (e.g., taboo expressions) coexist with "dynamic evolution" (e.g., modern interpretations of customs), this system establishes a dual-layer constraint mechanism of "rule filtering and semantic understanding." It offers a solution to cultural security challenges in ethnic minority knowledge graph construction [15], achieving both "rigid bottom-line protection" and "flexible dynamic adaptation".

The static constraint layer employs rule-based rapid filtering to cover "non-negotiable rigid boundaries" in cultural knowledge, such as intercepting queries about sensitive religious sites. This highly stable knowledge can be explicitly enumerated and solidified into rules to achieve millisecond-level interception, preventing misjudgments caused by semantic understanding deviations and reinforcing cultural security foundations.

The dynamic constraint layer utilizes an LLM-based semantic understanding module to address "flexible boundaries" in cultural knowledge that evolve over time, such as assessing the appropriateness of emerging expressions like "checking in at scripture turning" used by young tourists. By learning from the latest cultural cases—including user feedback from tourism platforms and guidelines from cultural institutions—the LLM comprehends

the relationship between expression context and cultural respect, resolving the limitation of traditional rule systems in adapting to cultural evolution.

Third, the integration of knowledge graphs and naked-eye 3D rendering technology reconstructs the positioning of cultural knowledge graphs as "cognitive engines." This integration lays a theoretical foundation for "knowledge-driven immersive experiences," extending the application of cultural knowledge graphs beyond traditional question-answering and recommendation systems [25]. This work redefines knowledge graphs from mere "information repositories" to "cognitive engines" for immersive technologies, representing a paradigm shift in digital cultural heritage preservation.

The paradigm shift operates as follows: whereas traditional "display-driven" 3D content relies on pre-rendered static assets that users passively watch, RLT2C decomposes cultural semantics into structured "entity-relationship-attribute" data. This enables the 3D rendering engine to dynamically generate content based on real-time queries—for example, when a user queries "Monlam Prayer Festival," the system not only retrieves textual information but also renders contextually appropriate 3D scenes based on the festival's association with Samye Monastery and historical crowd flow data within the knowledge graph.

This "knowledge-driven" paradigm enables an upgrade from "morphological recording" to "contextual restoration." By structuring cultural semantics into queryable and verifiable graphs, RLT2C facilitates precise and secure orchestration of multi-modal sensory outputs—including 3D visuals and scene sound effects—thereby ensuring the authenticity and integrity of intangible cultural heritage in digital formats.The high-precision 3D salient object detection and texture reconstruction methods in integral imaging [34] provide technical support for such immersive 3D rendering.

5.2. Practical Applications

RLT2C has implemented three core applications in Tibetan cultural tourism scenarios and possesses high concurrency processing capabilities, achieving the synergy between technological innovation and cultural respect, with measurable practical benefits:

5.2.1. Core Application Scenarios and Quantitative Results

Real-time knowledge update for tourism platforms: Through text-to-Cypher, authoritative cultural information (e.g., ceremony times, scenic spot introductions) is automatically converted into KG data. Compared with traditional manual maintenance, this reduces manual maintenance costs by 65% while avoiding deviations in cultural expressions caused by manual editing.

Culturally accurate naked-eye 3D guide systems: Currently deployed in three scenic spots in Lhasa, each of which has 92% positive user feedback. Unlike conventional audio guides, this system provides an "intuitive, spatially contextualized understanding" of cultural sites; for example, when a user queries "how to use a scripture-turning cylinder", the 3D model synchronously demonstrates the correct rotation direction and labels "the cultural meaning of clockwise rotation".

Intelligent content moderation for tourism UGC: Automatically identifies and blocks 98.7% of culturally inappropriate content. Relying on the dynamic-static cultural constraint system, it ensures the cultural compliance of platform content and avoids cultural misinterpretation.

5.2.2. Typical Application Cases

Immersive cultural scene restoration: When a user asks about the "Monlam Prayer Festival", the system not only retrieves textual information but also renders a 3D scene of Samye Monastery on that specific day, simulating the historical atmosphere and crowd

flow based on the graph's relational data. This "see-what-you-ask" capability dramatically enhances user engagement and comprehension.

Cultural safety mechanism verification: During a pilot test, a query attempting to access a 3D model of a sensitive ritual site was automatically intercepted by the rule-based static constraint system. This demonstrates the real-world effectiveness of the dual-safeguard approach in preventing cultural misrepresentation, ensuring that technological innovation in tourism does not come at the expense of cultural respect.

5.3. Limitations and Future Work

5.3.1. Core Limitations of the Current System

Insufficient ability to process colloquial text: Performance degrades on highly colloquial text with slang terms (e.g., social media posts like "Is the Samye Monastery ceremony worth attending?"). This is due to two factors: lightweight models cannot interpret the semantics of vague expressions, and LLMs lack training on colloquial corpora specific to Tibetan cultural tourism, leading to semantic mapping deviations.

Lag in cultural constraint system updates: As cultural norms evolve (e.g., modern adaptations of traditional rituals), the current system relies on manual updates of rules, resulting in "maintenance lag" and "subjectivity of expert judgments", making it difficult to capture cultural dynamics in real time.

Lack of support for multi-modal inputs: Currently, only text inputs are supported, and the system cannot respond to "image + text" queries (e.g., a user uploading a photo of a monastery and asking "What is the name of the ceremony held here"). This is because the KG has not yet integrated the association between image features (e.g., architectural contours) and textual knowledge.

5.3.2. Future Research and Optimization Directions

To address the above limitations and expand the system's application value, future work will focus on five key directions:

Improving colloquial text understanding: Incorporate contrastive learning to construct a colloquial corpus for Tibetan cultural tourism, including mapping pairs of "colloquial expression-standard semantics-cultural meaning". Enable the model to learn differences between similar expressions (e.g., "Is it worth attending" in the context of a ceremony means "whether to participate") and integrate a user feedback correction mechanism to continuously optimize parsing accuracy.

Building a crowdsourcing ecosystem for cultural rule updates: Design a three-level mechanism of "user contribution–expert review–system adaptation." Users submit new cultural cases (e.g., emerging expressions, custom changes), cultural experts review case validity, and the system automatically converts approved cases into constraint rules. LLMs will learn the logic behind these rules to realize dynamic updates of cultural constraints.

Extending multi-modal input processing: Construct a multi-modal cultural KG integrating "text + image + 3D model", and integrate image feature extraction technology (e.g., temple architectural contour recognition). Realize a closed loop of "image input → feature matching → KG entity positioning → multi-modal response" to support richer knowledge extraction scenarios.

Exploring in-depth integration with AR/VR technologies: Draw on the participatory activity design concept proposed by Silva et al. [35] to integrate RLT2C with AR (e.g., users scanning temple buildings with AR to get real-time labels of architectural historical purposes based on the KG) and VR (e.g., creating "digital twin" replicas of Tibetan cultural heritage sites for remote education, cultural preservation, and virtual pilgrimage). The

use of reflective polarizers for compact, high-performance AR 3D displays [36] offers a hardware foundation for such integrations.

Optimizing 3D content generation with PCG technology: Leverage the KG's semantic rule base (e.g., "a ceremony scene must include prayer flags and scripture-chanting crowds") to drive procedural content generation (PCG) for 3D scenes. This allows the system to dynamically create historically plausible architectural details or ritual animations, further enhancing the richness and adaptability of immersive experiences.The tunable and adaptive salient perception network for integral imaging 3D salient object detection [37] provides a reference for achieving high-accuracy multi-target segmentation in such dynamically generated 3D scenes.

6. Conclusions

This paper introduces RLT2C (Rule+LLM-Verify for Text-to-Cypher), a novel methodology that integrates rule-based guidance with large language model (LLM) verification for automated knowledge graph construction in Tibetan cultural tourism. Experimental evaluation on 1200 Tibetan tourism-related texts demonstrates that RLT2C outperforms all baseline methods across four key metrics: construction efficiency (14.5 triples/100 words), relationship accuracy (91.5%), local semantic adaptability (87.9%), and graph redundancy rate (5.4%).

The framework incorporates key innovations, including lightweight-large model collaboration, dynamic-static cultural constraint integration, and Neo4j full-link optimization, which collectively address persistent challenges in Tibetan tourism KG development. These challenges include high manual dependency, inadequate semantic adaptability, and sluggish dynamic response. Through its dual-safeguard mechanism, the system achieves 98.1% cultural compliance, making it particularly suitable for culturally sensitive domains.

RLT2C establishes an extensible technical paradigm for digitally preserving and presenting ethnic minority cultures. Its structured knowledge graph foundation specifically supports knowledge-driven naked-eye 3D immersive experiences, while the modular architecture facilitates straightforward adaptation to diverse cultural domains beyond Tibetan tourism. This approach paves the way for developing more accurate, efficient, and culturally respectful digital knowledge systems worldwide.

Author Contributions: Conceptualization, K.W.; Data curation, K.W., S.Y. (Shuai Yan), Z.L., X.Y., F.L. and B.J.; Formal analysis, K.W. and S.Y. (Shuai Yan); Funding acquisition, K.W.; Investigation, K.W., S.Y. (Shuai Yan), Z.L., X.Y., F.L. and B.J.; Methodology, K.W.; Project administration, K.W. and S.Y. (Shuai Yan); Resources, K.W. and H.D.; Software, K.W. and S.Y. (Shengying Yang); Supervision, K.W., S.Y. (Shengying Yang) and H.D.; Validation, K.W., Z.L., X.Y., F.L., B.J. and S.Y. (Shengying Yang); Visualization, K.W. and S.Y. (Shuai Yan); Writing—original draft, K.W.; Writing—review & editing, K.W., S.Y. (Shuai Yan), Z.L., X.Y., S.Y. (Shengying Yang) and H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Science and Technology Projects of Xizang Autonomous Region, China, project title "Research and Application of LLM-based Intelligent Tourism Service System for Xizang", grant number XZ202401ZY0008.

Data Availability Statement: The datasets presented in this article are not readily available due to the following reasons: Ongoing Research Constraints: The data are an integral part of a currently active study, and premature disclosure could compromise the integrity of future publications and project outcomes. Cultural and Political Sensitivities: Tibetan cultural tourism data involve unique regional characteristics, and their collection is subject to strict cultural preservation protocols and political regulations, making public distribution challenging. Requests for limited access to the datasets under specific conditions (e.g., academic collaboration or non-commercial research) may be directed to

the corresponding author via email at yangshengying@cdjcc.edu.cn. Each request will be evaluated based on compliance with cultural security guidelines and project progress.

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

References

1. Hogan, A.; Blomqvist, E.; Cochez, M.; D'amato, D.; De Melo, G.; Gutierrez, C.; Kirrane, S.; Labra Gayo, J.E.; Navigli,B.; Neumaier, S. Knowledge Graphs. *ACM Comput. Surv.* **2022**, *54*, 71. [CrossRef]
2. Lehmann, J.; Isele, R.; Jakob, M.; Jentzsch, A.; Kontokostas, D.; Mendes, P.N.; Hellmann, S.; Morsey, M.; Van Kleef, P.; Auer, S.; et al. DBpedia—A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semant. Web* **2015**, *6*, 167–195. [CrossRef]
3. Zhou, W. Design and Implementation of Personalized Tourism Recommendation System on Basis of Knowledge Graph. In Proceedings of the 2024 3rd International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI), Sanya, China, 24–26 January 2024; pp. 64–68. [CrossRef]
4. Steinigen, D.; Teucher, R.; Ruland, T.H.; Rudat, M.; Flores-Herr, N.; Fischer, P.; Milosevic, N.; Schymura, C.; Ziletti, A. Fact Finder—Enhancing Domain Expertise of Large Language Models by Incorporating Knowledge Graphs. *arXiv* **2024**, arXiv:2408.03010. [CrossRef]
5. Xiao, D.; Wang, N.; Yu, J.; Zhang, C.; Wu, J. A Practice of Tourism Knowledge Graph Construction based on Heterogeneous Information. In Proceedings of the 19th Chinese National Conference on Computational Linguistics, Haikou, China, 30 October–1 November 2020; Chinese Information Processing Society of China: Haikou, China, 2020; pp. 939–949. Available online: https://aclanthology.org/2020.ccl-1.87/ (accessed on 15 October 2025).
6. Ren, X.; Tang, J.; Yin, D.; Chawla, N.; Huang, C.A Survey of Large Language Models for Graphs. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24), Barcelona, Spain, 25–29 August 2024; pp. 6616–6626. [CrossRef]
7. Szárnyas, G.; Püroja, D.; Boncz, P.; Bebee, B.; Gosnell, D.; Birler, A.; Deutsch, A.; Wu, M.; Fletcher, G.; Gabb, H.A.; et al. The Linked Data Benchmark Council (LDBC): Driving Competition and Collaboration in the Graph Data Management Space. In *Performance Evaluation and Benchmarking, Proceedings of the TPCTC 2023, Vancouver, BC, Canada, 28 August–1 September 2023*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2024; Volume 14247, p. 7. [CrossRef]
8. Cardoso Cardoso, R.C.; Sohn, A.P.L.; Ferasso, M.; Júnior, S.P. Open Innovation in the Tourism Field: A Systematic Literature Review. *J. Open Innov. Technol. Mark. Complex.* **2024**, *10*, 00359. [CrossRef]
9. Sun, H.K. Historical Relics in Sino-Tibetan Languages from the Perspective of Cognate Numerals. *Lang. Sci.* **2018**, *17*, 561–579. [CrossRef]
10. Fan, Z.; Chen, C. CuPe-KG: Cultural perspective–based knowledge graph construction of tourism resources via pretrained language models. *Inf. Process. Manag.* **2024**, *61*, 103646. [CrossRef]
11. Xue, S.; Jiang, C.; Shi, W.; Cheng, F.; Chen, K.; Yang, H.; Zhang, Z.; He, J.; Zhang, H.; Wei, G.; et al. DB-GPT: Empowering Database Interactions with Private Large Language Models. *arXiv* **2024**, arXiv:2312.17449. [CrossRef]
12. Yang, C.; Li, C.; Hu, X.; Yu, H.; Lu, J. Enhancing Knowledge Graph Interactions: A Comprehensive Text-to-Cypher Pipeline with Large Language Models. *Inf. Process. Manag.* **2026**, *63*, 104280. [CrossRef]
13. El Boujddaini, F.; Laguidi, A.; Mejdoub, Y.A Survey on Text-to-SQL Parsing: From Rule-Based Foundations to Large Language Models. In Proceeding of the International Conference on Connected Objects and Artificial Intelligence (COCIA2024), Casablanca, Morocco, 8–10 May 2024; Mejdoub, Y., Elamri, A., Eds.; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2024; Volume 1123, pp. 41–52. [CrossRef]
14. Zhang, K.; Lin, X.; Wang, Y.; Zhang, X.; Sun, F.; Cen, J.; Tan, H.; Jiang, X.; Shen, H. ReFSQL: A Retrieval-Augmentation Framework for Text-to-SQL Generation. In Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2023, Singapore, 6–10 December 2023; pp. 664–673. [CrossRef]
15. Sun, G.; Shen, R.; Jin, L.; Wang, Y.; Xu, S.; Chen, J.; Jiang, W. Instruction Tuning Text-to-SQL with Large Language Models in the Power Grid Domain. In Proceedings of the 2023 4th International Conference on Control, Robotics and Intelligent System, Guangzhou, China, 25–27 August 2023; pp. 59–63. [CrossRef]
16. Yao, L.; Peng, J.; Mao, C.; Luo, Y. Exploring Large Language Models for Knowledge Graph Completion. In Proceedings of the 2025 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2025), Cape Town, South Africa, 6–11 April 2025. [CrossRef]
17. Liu, H.; Wang, S.; Zhu, Y.; Dong, Y.; Li, J. Knowledge Graph-Enhanced Large Language Models via Path Selection. In Proceedings of the Findings of the Association for Computational Linguistics: ACL 2024, Bangkok, Thailand, 11–16 August 2024; pp. 6311–6321. [CrossRef]

18. He, W.; Xu, Y.; Yu, Q. BERT-BiLSTM-CRF Chinese Resume Named Entity Recognition Combining Attention Mechanisms. In Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering, Dalian, China, 17–19 November 2023; pp. 542–547. [CrossRef]
19. Wang, Y.; Liu, J.; Wang, W.; Chen, J.; Yang, X.; Sang, L.; Wen, Z.; Peng, Q. Construction of Cultural Heritage Knowledge Graph Based on Graph Attention Neural Network. *Appl. Sci.* **2024**, *14*, 8231. [CrossRef]
20. Gao, J.; Peng, P.; Lu, F.; Claramunt, C.; Qiu, P.; Xu, Y.Mining Tourist Preferences and Decision Support via Tourism-Oriented Knowledge Graph. *Inf. Process. Manag.* **2023**, *61*, 103523. [CrossRef]
21. Luo, W.; Dang, H.; Liu, W.; Gao, Y. Research on the Construction of a Knowledge Graph for Intangible Cultural Heritage in Tibet Based on Big Data. *Tibet Sci. Technol.* **2022**, *1*, 75–80.
22. Becker, J.; Botsch, M.; Cimiano, P.; Derksen, M.; Elahi, M.; Maier, A.; Maile, M.; Pätzold, I.; Penningroth, J.; Reglin, B.; et al. Virtual Reality Based Access to Knowledge Graphs for History Research. In *Semantic Systems. The Power of AI and Knowledge Graphs*; Pellegrini, T., Ed.; IOS Press: Amsterdam, The Netherlands, 2023; pp. 143–157. [CrossRef]
23. Yang, S.; Hou, M. Knowledge Graph Representation Method for Semantic 3D Modeling of Chinese Grottoes. *Herit. Sci.* **2023**, *11*, 266. [CrossRef]
24. Wang, J.; Zakaria, S.A. Design Application and Evolution of 3D Visualization Technology in Architectural Heritage Conservation: A CiteSpace-Based Knowledge Mapping and Systematic Review (2005–2024). *Buildings* **2025**, *15*, 1854. [CrossRef]
25. Nguyen, C.; Le, M.T.; Yoon, D.-I.; Kim, H.-K. 3D Graphics Visualization and Context Information Service for a Virtual Tourist System. *J. Ubiquitous Converg. Technol.* **2007**, *1*, 47–52.
26. Zhang, X.; Zhang, P.; Luo, S.; Tang, J.; Wan, Y.; Yang, B.; Huang, F. CultureSynth: A Hierarchical Taxonomy-Guided and Retrieval-Augmented Framework for Cultural Question-Answer Synthesis. *arXiv* **2025**. [CrossRef]
27. Wang, X.; Gao, X.; Fu, Z.; Chen, X.; Wang, X. Design and Implementation of Event Knowledge Graph Construction Platform Based on Neo4j. In Proceedings of the 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballar, India, 29–30 April 2023; pp. 1–6. [CrossRef]
28. Shi, L.; Tang, Z.; Zhang, N.; Zhang, X.; Yang, Z. A Survey on Employing Large Language Models for Text-to-SQL Tasks. *ACM Comput. Surv.* **2025**, *58*, 54, 1–37. [CrossRef]
29. Dai, Y.; Wang, S.; Xiong, N.N.; Guo, W. A Survey on Knowledge Graph Embedding: Approaches, Applications and Benchmarks. *Electronics* **2020**, *9*, 750. [CrossRef]
30. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [CrossRef]
31. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33*, 159–174. [CrossRef]
32. Gao, Y.; Liu, Y.; Li, X.; Shi, X.; Zhu, Y.; Wang, Y.; Li, S.; Li, W.; Hong, Y.; Luo, Z.; et al. A Preview of XiYan-SQL: A Multi-Generator Ensemble Framework for Text-to-SQL. *arXiv* **2024**, arXiv:2411.08599. [CrossRef]
33. Ibrahim, N.; Aboulela, S.; Ibrahim, A.; Kashef, R.A Survey on Augmenting Knowledge Graphs (KGs) with Large Language Models (LLMs): Models, Evaluation Metrics, Benchmarks, and Challenges. *Discov. Artif. Intell.* **2024**, *4*, 76. [CrossRef]
34. Li, Q.; Li, W.; Zhao, D.; Dong, B.; Kou, Y.; Li, X.; Wang, X. High-precision integral imaging 3D salient object detection and reconstruction with texture features based on E2E-TransGAN. *Opt. Express* **2024**, *32*, 36329–36343. [CrossRef] [PubMed]
35. Silva, C.; Zagalo, N.; Vairinhos, M.Towards Participatory Activities with Augmented Reality for Cultural Heritage: A Literature Review. *Comput. Educ. X Real.* **2023**, *3*, 100044. [CrossRef]
36. Li, Q.; He, W.; Deng, H.; Zhong, F.Y.; Chen, Y. High-performance reflection-type augmented reality 3D display using a reflective polarizer. *Opt. Express* **2021**, *29*, 9446–9453. [CrossRef] [PubMed]
37. Li, Q.; Zhao, D.; Li, W.-Y.; Dong, B.-Z.; Li, X.-W.; Wang, X.-R. Adaptive Topology-Driven Integral Imaging 3D Salient Object Detection for Complex Multi-Target Scenes. *Opt. Laser Technol.* **2025**, *192*, 113684. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.