

www.ksii.or.kr

2026년 춘계학술발표대회

2026년 5월 7일(목)~9일(토)

제주 휴닉스아일랜드

주최 및 주관

사단법인 한국인터넷정보학회

후원

(주)워드마이너, (주)나루씨큐리티, (주)이글루코퍼레이션, 국방전산정보원, 한국AI사물인터넷협회,
(주)아이티센텍, 국방혁신기술보안협회, 한화시스템(주), (주)LIG시스템



www.ksii.or.kr



KOREAN SOCIETY FOR INTERNET INFORMATION
사단법인 한국인터넷정보학회



논문발표

5월 8일(금) 9:00~10:30 (Room 1 / 스톤홀1)

세션 A-4

좌장: 한재호 교수(고려대)

인공지능 4

- A-4-1 비디오 확산 모델의 입력 조건 변형을 통한 로봇 조작 데이터 증강
정준호, 문창욱, Dong Jin, 구영현(세종대)
- A-4-2 다중 작업 EyeCLIP 기법을 이용한 안저 영상 기반 전신 바이오마커 예측
Muhammad Omer, Le Thi Kim Ngoc, 추현승(성균관대)
- A-4-3 VLM 기반 병해충 이미지 자동 필터링 프레임워크
최준희, 구영현, 윤학림, 이지민(세종대)
- A-4-4 실환경 강화학습을 위한 All-action Planning 기반 Prioritized Dyna-DQN
허진욱, 권용진(한국항공대)
- A-4-5 신경망 구조 탐색 기반 FCOS 및 피라미드 비전 트랜스포머를 이용한 작물 해충 진단
Huanyu Zhou, Dong Jin, Helin Yin, 구영현(세종대)
- A-4-6 non-IID 데이터 환경에서 연합 언어 모델 학습의 최악 도메인 강건성 향상을 위한 어텐션 기반 집계 모델링
Ivanova Daniela, LI YALONG, 조민호(고려대)

5월 8일(금) 9:00~10:30 (Room 2 / 스톤홀2)

세션 F

좌장 : 김문성 교수(서울신학대)

네트워크 & 메타버스

- F-1 시공간적 네트워크 트래픽 예측을 위한 프롬프트 엔지니어링 기반 대규모 언어 모델
Thi Le Quyen Nguyen, Yifei Tian, 추현승(성균관대)
- F-2 웹 QoE 환경에서 사용자 이탈의 임계 기반 해석
이승현(국방대), 전병진(육군사관학교), 임효영(국방대)
- F-3 네트워킹 기술 사양에서 RAG를 위한 문서 청킹 전략
Sardar Jaffar Ali, Yifei Tian(성균관대), 김문성(서울신학대), 추현승(성균관대)
- F-4 차세대 통신에서 대규모 언어 모델(LLMs)의 진화
보반비, Yifei Tian, 추현승(성균관대)
- F-5 가상 전사관 내 RAG 기반 AI 도슨트의 데이터 및 지식 재사용성 구현 사례
김인용, 조형구, 공지훈(홍익대), 김인권(아이스큐브랩), 김장환, 서채연, 김영철(홍익대)

가상 전시관 내 RAG 기반 AI 도슨트의 데이터 및 지식 재사용성 구현 사례

김인용¹, 조형구², 공지훈³, 김인권⁴, 김장환⁵, 서채연⁶, 김영철^{7*}
 홍익대학교 소프트웨어융합학과^{1,3,5,6,7*}, 홍익대학교 게임학부²
 아이스크류브랩⁴

e-mail : citefred@gmail.com¹, whgudrn123@gmail.com², go400s@gmail.com³,
 ingunbi@gmail.com⁴, {lentoconstante⁵, chaeyun⁶, bob^{7*}}@hongik.ac.kr

Implemented Practice on Reusing Data and Knowledge of RAG-driven AI Docent in Virtual Exhibitions

Inyong Kim¹, HyeongGu Jo², Jihoon Kong³, Inkwon Kim⁴, Janghwan Kim⁵,
 Chaeyun Seo⁶, R. Young Chul Kim^{7*}

1. 서론

가상현실(VR) 전시는 시공간적 제약을 극복하는 대안이나, 대다수 일회성 하드코딩 방식으로 제작되어 전시 종료 후 콘텐츠 재사용이 어렵다. 선행 연구에서는 이를 해결하고자 건축 도면(BIM) 기반의 공간 구현과 웹 CMS(Content Management System) 연동을 통한 시각적 데이터 재사용성을 확보했고, 문제는 도슨트의 AI 적용이 미비했다[1], [2].

최근 몰입감 향상을 위해 도입되는 AI 도슨트는 프롬프트 고정이나 미세조정(Fine-tuning) 방식 사용 시 신규 전시마다 발생하는 재학습 비용과 환각(Hallucination) 현상이 한계로 지적된다. 이를 해결하고자 기존 아키텍처에 검색 증강 생성(RAG) 파이프라인을 결합하여 AI 도슨트의 지식(Knowledge) 재사용 방안을 제안한다. 본 연구는 2장 관련 연구, 3장 설계, 4장 구현 및 검증, 5장 결론으로 구성된다.

2. 관련 연구

가상 전시 환경의 지속 가능성 확보를 위해, 선행 연구[1], [2]에서는 웹 CMS 연동을 통해 소스 코드 수정 없이 전시 작품을 교체하는 시각적 데이터 재사용성 모델을 구축한 바 있다. 시각적 경험을 넘어서는 지능형 인터랙션 측면에서, AI 도슨트의 역량이 관람객 만족도에 미치는 긍정적 영향이 실증되었다[3].

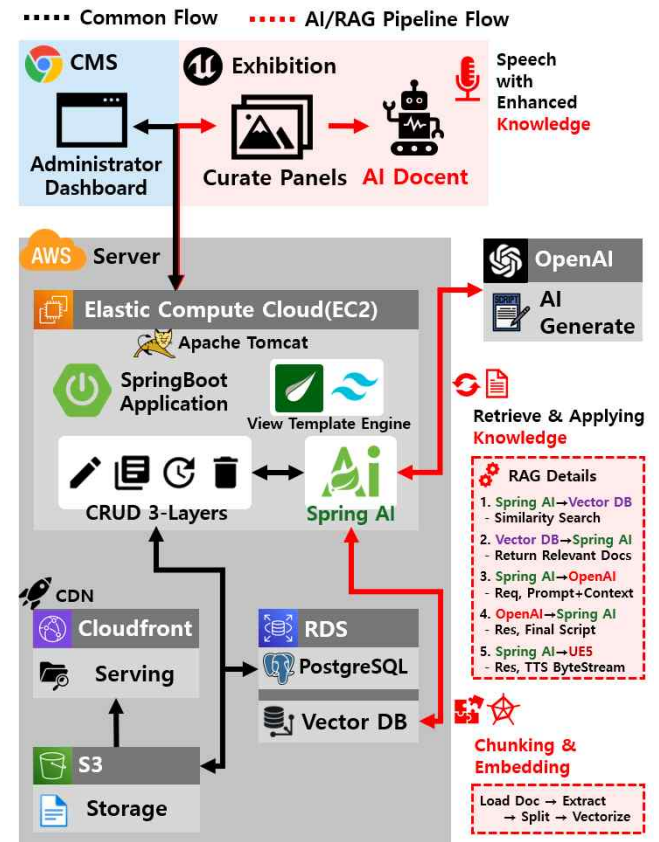
하지만, 기존의 LLM 미세조정 방식은 신규 전시 기획 때마다 발생하는 재학습 비용으로 인해 운영 효율성을 저해한다[4]. 최근 멀티모달 RAG 기술로 지식 갱신의 유연성을 확보하려는 시도가 이루어졌으나[5], 전시 기획자가 시각 데이터와 도메인 지식(Vector DB)을 통합 관리할 수 있는 파이프라인에 대한 고찰은 여전히 부족하다. 따라서 본 연구는 CMS 구조에 RAG 시스템을 결합하여, 운영 편의성과 AI 도슨트의 지식 재사용성을 동시에 달성하는 통합 파이프라인을 제시한다.

3. 지식 및 데이터 재사용성 기반 파이프라인 설계

본 연구가 제안하는 프레임워크는 선행 연구[1], [2]에서 구축된 웹 기반 CMS와 3D 가상 전시 공간을 기반으로, AI 도슨트의 지식을 동기화하는 검색 증강 생성(RAG) 파이프라인을 결합한 구조이다.

기존 시스템이 애플리케이션 재배포 없는 시각 데이터의 동적 로드(Dynamic Load)에 그쳤다면, 본 시스템은 그림 1과

같이 관리자가 CMS에 신규 전시 이미지와 함께 문서(PDF)나 작가 노트를 업로드하는 즉시 AI의 도메인 지식까지 단일 파이프라인 내에서 갱신되도록 설계되었다.



(그림 1) 시스템 전체 구성도

이를 위해 본 연구는 Spring AI 기반 RAG 파이프라인을 구축하여 지식 갱신 비용을 최소화하였다. 시스템은 업로드된 PDF 포맷 등 문서 파일에서 텍스트를 추출, 청킹(Chunking) 및 임베딩하여 Vector DB(pgvector)에 저장한다. 이후 관람객이 가상 전시관 내 특정 작품에 접근하면, RAG 파이프라인이 Vector DB에서 최적의 문서를 검색(Retrieval)하고 LLM 프롬프트에 주입하여 맞춤형 해설을 생성한다.

표 1은 제안 시스템에서 시각 데이터와 지식이 통합적으로 업데이트되어 음성으로 반환되기까지의 동기화 흐름을 요약

한 것이다.

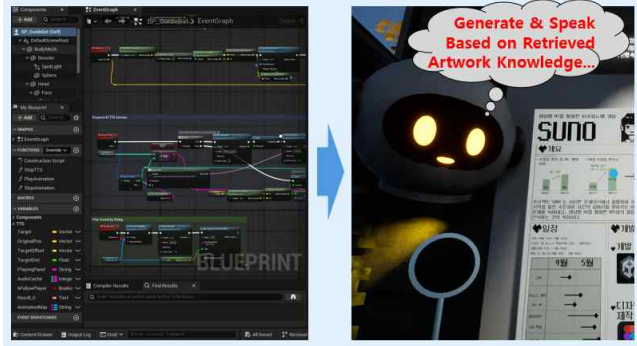
(표 1) 데이터 및 지식 업데이트 프로세스

단계	구분	역할
1단계	CMS업로드	관리자가 웹 인터페이스를 통해 신규 전시 이미지 및 문서(PDF) 업로드
2단계	지식 자동화	백엔드에서 텍스트 추출, 청킹 및 임베딩 후 Vector DB(pgvector) 즉시 갱신
3단계	데이터 로드	Unreal Engine(UE) 클라이언트 실행 또는 레벨 이동 시 API를 통해 최신 전시 메타데이터 수신
4단계	해설 생성	관람객 특정 작품 접근 시 RAG 파이프라인이 갱신된 지식을 자동 검색하여 맞춤형 해설 생성
5단계	음성 변환	응답된 해설을 자연스러운 AI 보이스 바이츠 트립으로 변환 및 응답, UE 클라이언트에서 재생

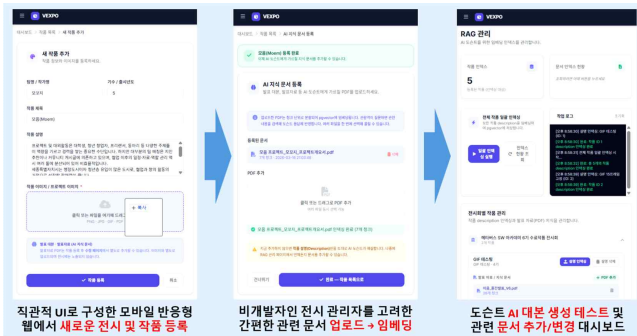
4. 구현 및 검증

4.1 구현 환경 및 시스템 구성

통합 CMS 및 백엔드는 AWS 환경의 Spring Boot로 구축되었다. RAG 지식베이스는 Spring AI와 PostgreSQL의 pgvector를 활용하여 벡터 검색을 처리한다. 그림 2는 가상 전시관(UE5) 내 AI 도슨트의 상호작용 및 지식 렌더링 구현 결과이다. 그림 3은 비개발자 관리자로 손쉽게 지식을 갱신할 수 있는 플랫폼의 모습이다.



(그림 2) AI 도슨트 인터랙션 로직 Blueprint 구현 및 렌더링 결과

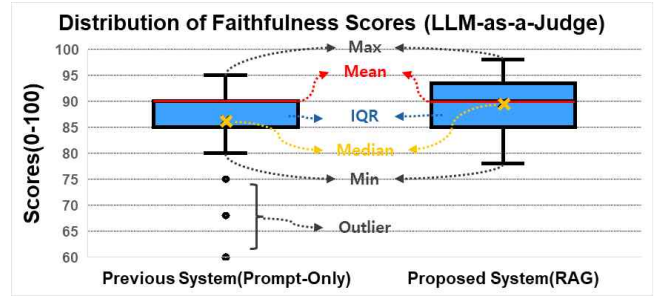


(그림 3) 전시 관련 정보 업로드, 임베딩, 동작 테스트 UI 구성

4.2 AI 도슨트 지식 재사용성 및 신뢰성 검증

도슨트 해설의 신뢰성을 검증하기 위해 총 50회의 대본을 자동 생성하여 정량 평가를 수행하였다. 단순 텍스트 일치율의 한계를 극복하고자 'LLM-as-a-Judge' 기법을 도입하여, 원본 문서(PDF)의 심층 지식 반영도를 평가하였다.

평가 결과, 기존 프롬프트 방식은 평균 86.3점이나 간헐적 환각으로 최저 60점(표준편차 8.3)까지 하락하는 불안정성을 보였다. 반면, 제안 시스템은 Vector DB 동적 검색을 통해 평균 89.6점을 달성하였고, 그림 4와 같이 최저 78점을 방어(표준편차 4.6)하며 지식 안정성을 확보하였다. 이는 추가적인 모델 재학습 없이도 CMS에 업로드된 문서만으로 고품질의 맞춤형 해설을 일정하게 제공할 수 있음을 실증한다.



(그림 4) 지식 반영 방식에 따른 사실 부합률 분포 비교

생성 대본의 평균 정보량이 1.3배(668→859자) 증가해 심층 지식의 대본 반영 상태를 확인 할 수 있었다. 표 2는 두 방식의 비교 지표이다.

(표 2) 지식 반영 방식에 따른 성능 및 운영 효율 비교

구분	프롬프트 엔지니어링(기존)	RAG 기반 동적 검색(제안)
지식 부합률	평균 86.3점	평균 89.6점
대본 길이	평균 668자(기초 요약)	평균 859자(정보 1.3배 ↑)
참조 시간	측정 불가(해당 기능 없음)	평균 291ms
응답 시간	평균 5.75초	평균 8.33초
재사용성	낮음(변경 재입력 필요)	높음(문서 업로드 시 갱신)

5. 결 론

본 연구는 통합 CMS와 RAG 파이프라인을 결합하여 가상 전시관 AI 도슨트의 지식 재사용성을 성공적으로 구현하였다. 관리자는 모델 재학습 없이 즉각적인 지식 갱신이 가능해 전시 운영 및 유지보수 비용을 획기적으로 절감할 수 있다.

향후 연구에서는 응답 지연 시간(Latency)을 은닉하는 비동기 UX 마스킹 기법을 적용할 예정이다. 나아가 AI가 전시 기획 의도와 관람객 성향을 분석하여, 가상 공간의 테마를 동적으로 변환하는 지능형 메타버스 프레임워크로 본 연구를 확장할 계획이다.

Acknowledgment

본 연구는 2025년도 문화체육 관광부의 재원으로 한국콘텐츠진흥원(과제명: 인공지능 기반 대화형 멀티모달 인터랙티브 스토리텔링 3D장면 저작 기술 개발, 과제번호:RS-2023-00227917, 기여율:100%) 지원과 한국연구재단의 4단계 두뇌한국21사업(과제명: 초분산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호:202003520007)의 지원을 받아 수행된 연구임. 또한, 홍익대학교 메타버스 융합 SW 아카데미 6기 교육생들의 프로젝트 결과물이다.

참 고 문 헌

- [1] 김인용, 조형구, 마예은, 오재은, 공지훈, 김장환, 김영철, “건축 도면 기반 VR 가상 전시관 개발 모범 사례,” 2025 한국인터넷정보학회 춘계학술발표대회, 2025.
- [2] 김인용, 조형구, 마예은, 공지훈, 김기두, 김장환, 김영철, “통합 콘텐츠 관리 시스템 기반 재사용 VR 전시관 구현 사례,” 2025 한국스마트미디어학회 춘계학술대회, 2025.
- [3] 강지성, 신헌덕, “도슨트 역량에 대한 평가가 전시 관람의도에 미치는 영향: 인공지능 로봇 도슨트의 효과를 중심으로,” 한국콘텐츠학회논문지, 제24권, 제12호, pp.509-516, 2024.
- [4] 김지민, 김누리, “도메인 특화 AI 도슨트 구현 및 성능 비교 연구: RAG와 파인튜닝 모델을 중심으로,” in Proc. of KIIT Conference, 제주, 2025.
- [5] 강채원, 김예민, 유채민, 윤소은, 이윤서, 유건아, “멀티모달 검색-증강 생성과 대규모 언어 모델을 활용한 맞춤형 미술관 도슨트 시스템,” 디지털콘텐츠학회논문지, 제27권, 제2호, pp.473-484, 2026.