

한국 전통문양의 생성형 AI 학습 데이터 구축과 활용 사례 및 전략 연구

A Study on Establishment and Utilization of Generative AI Training Data on Korean Traditional Patterns

심정택(주저자) · 정재훈(교신저자)

Jungtaek Shim(First Author) · Jaehoon Jeong(Corresponding Author)

서울시립대학교/국가유산진흥원 University of Seoul/Korea Heritage Agency(icanjit@uos.ac.kr)

서울시립대학교 University of Seoul(jaehoon@uos.ac.kr)

본 연구는 초거대 인공지능(Artificial Intelligence, AI) 시대에 대응하여 한국 전통문양의 디지털 자산화와 문화콘텐츠 산업 활용을 위한 생성형 AI 학습 데이터 구축 사례를 분석하고, 그 활용 전략을 제시하는 것을 목적으로 한다. 2024년 국가유산진흥원이 수행한 「한국 전통문양 생성형 AI 학습 데이터 구축 사업」을 탐색적 사례 연구(Exploratory Case Study)로 선정하여, 해당 사례의 데이터 수집, 정제, 가공, 품질관리 등 전 과정을 분석하고, 실무적 성과를 구조화하여 구체적인 방법론과 활용 가능성을 제시하였다. 특히 전통문양의 형태별 · 용도별 · 시대별 분류체계와 한 · 영 이미지 캡셔닝 방식의 적용 사례를 고찰하고, 게임, 메타버스, 디지털 아트, AR/VR 등 문화콘텐츠 산업 분야의 적용 가능성과 공공데이터 개방 및 생태계 조성 전략을 제안하였다. 이러한 분석은 한국 전통문양의 문화적 정체성과 독창성을 보존하면서 디지털 문화유산의 활성화 및 산업적 활용 가능성을 제고하는 데 실질적으로 기여할 것이다.

주제어: 초거대 인공지능, 전통문양 데이터, 생성형 AI, 문화콘텐츠 산업, 데이터 생태계

This study examines the construction of generative AI training datasets for the digital assetization and industrial application of traditional Korean patterns within the contemporary era of hyper-scale artificial intelligence (AI). Utilizing the 2024 "Korean Traditional Pattern Data Establishment Project" by the Korea Heritage Agency as an exploratory case study, this research analyzes the end-to-end lifecycle of data development, including collection, refinement, processing, and quality control. By structuralizing these empirical outcomes, the study proposes concrete methodologies and versatile application frameworks for the digital cultural content industry. Specifically, the research evaluates classification systems categorized by morphology, historical utility, and era, alongside the implementation of bilingual (Korean-English) image captioning techniques for multimodal AI training. Furthermore, it explores the integration of these datasets into emerging sectors such as the metaverse, digital art, and Extended Reality (AR/VR/XR), while outlining strategic pathways for public data dissemination and the cultivation of a sustainable data ecosystem. This analysis provides an empirical foundation for promoting the digital revitalization of traditional cultural heritage while ensuring the preservation of its cultural identity and aesthetic authenticity.

Keyword: Hyper-scale AI, Traditional Pattern Data, Generative AI, Cultural Content Industry, Data Ecosystem

최초투고일: 2025. 06. 29

게재확정일: 2025. 10. 09

Copyright 2026 THE KOREAN ACADEMIC SOCIETY OF BUSINESS ADMINISTRATION

This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0, which permits unrestricted, distribution, and reproduction in any medium, provided the original work is properly cited.

I. 서론

21세기 들어 한류(韓流)는 대중음악(K-팝)과 드라마를 매개로 세계적 문화 현상으로 부상하였으며, 이를 통해 한국 대중문화가 전 세계로 확산되고 있다(한국국제교류재단, 2024). 이러한 확산은 한국 문화 전반에 대한 관심으로 확장되면서, 대중문화뿐 아니라 전통문화에 대한 수요와 탐구 역시 심화되고 있다.

최근 외국인들의 국립중앙박물관 방문이 급증하고 있으며, 이는 한류를 통해 형성된 대중문화에 대한 관심이 전통문화에 대한 관심과 체험으로 확장되고 있음을 보여준다(국립중앙박물관, 2024). 이러한 한류와 전통문화 간의 연결성은 전통문화 확산을 도모할 수 있는 전환점으로 작용하고 있으며, 전통문화의 디지털 자산화 및 인공지능 기술 활용을 통한 현대적 계승 방안의 필요성을 부각시키고 있다(정명희, 2023).

오늘날 초거대(Hyper-scale) 인공지능(Artificial Intelligence, AI)의 발전은 데이터 기반 자동 생성 기술을 통해 디지털 문화유산의 구축과 활용 방식에 구조적 변화를 야기하고 있다(허명숙&천면중, 2021). 특히 생성형 AI(Generative AI) 기술은 이미지 생성, 창작 과정 지원, 문화유산 복원 등 다양한 분야에서 활용도가 급증하고 있으며(정보통신정책연구원, 2025), 문화콘텐츠 산업 전반에 걸쳐 새로운 성장 동력으로 주목받고 있다(권구민&김현석, 2023).

한국 전통문양은 건축물, 공예품, 생활소품 등 전통문화 자산 전반에 걸쳐 활용되어 온 시각적 상징 체계로, 문화적 정체성과 고유성을 지닌 국가적 문화 자산이다. 우리 전통문화는 민족문화의 소산으로 생활관습, 정서, 종교, 신앙적 의미를 담고 있는 미의 식적 표현이며, 민족의 역사를 보여주는 중요한 문화

유산이다(권기형&김성남, 2011). 그러나 디지털 데이터 차원에서의 체계적 구축은 여전히 제한적이며, 특히 인공지능 학습을 위한 고품질 데이터의 확보는 충분히 이루어지지 못하고 있다(에이치씨아이플러스, 2025).

특히 글로벌 AI 이미지 생성 서비스에서 한국 전통 문양을 검색하면 중국, 일본 등 주변국 문화와 유사한 이미지가 생성되는 문제가 빈번히 발생하고 있으며, 이는 한국 전통문화의 정체성 훼손과 글로벌 콘텐츠 시장에서의 경쟁력 약화로 이어질 우려가 있다(국가유산진흥원, 2025). 이러한 현상은 생성형 AI의 편향성과 문화적 대표성 부족에서 기인하며, 문화유산 분야에서 AI 활용 시 발생할 수 있는 주요 위험 요소로 지적되고 있다(Foka and Griffin, 2024).

전통문양 데이터는 문화유산 활용과 더불어 문화콘텐츠 산업 전반에 광범위한 파급효과를 미치고 있다. 고품질 데이터 구축을 통해 게임, 메타버스, AR/VR(Augmented Reality/Virtual Reality), 디지털 아트 등의 산업 활용도를 제고할 수 있으며(노형후&박상명, 2023), 공공데이터 개방과 민간 서비스 확산을 통한 디지털 문화산업 생태계 조성에도 기여할 수 있다(서유덕, 2024). 전통문양은 단순한 장식 요소를 넘어 신성(神聖), 복(福), 수호, 번영 등 다양한 의미를 내포하고 있으며, 의례적 기능과 심미적 기능을 동시에 수행해 왔다(김혜진, 2019).

이러한 배경 하에 본 논문의 연구 목적은 초거대 인공지능 시대를 맞아 한국 전통문양을 생성형 AI 학습 데이터로 구축한 실증사례를 분석하고, 이를 기반으로 문화콘텐츠 산업, 공공서비스 및 교육 분야에서의 활용 방안과 지속가능한 데이터 생태계 조성 전략을 제시하는 데 있다.

본 연구는 탐색적 사례 연구(Exploratory Case Study)로, 2024년 국가유산진흥원이 수행한 「한국

전통문양 생성형 AI 학습 데이터 구축 사업」¹⁾(이하 '2024년 데이터 구축 사업')을 단일 사례로 선정하였다(George and Bennett, 2005; 국가유산진흥원, 2024). 이 사례는 생성형 AI와 전통문화의 융합이라는 정책적·산업적 선도 모델로, 데이터 자산화와 문화 주권에 관한 이론적 함의를 제공한다(George and Bennett, 2005). 본 논문에서는 이 사례의 데이터 수집, 정제, 가공, 품질관리 과정을 체계적으로 정리하고, 전통문양의 형태·용도·시대별 분류체계와 이미지 캡셔닝(Image Captioning) 기법 등을 상세히 고찰하였다(국가유산진흥원, 2025).

본 논문의 연구 질문은 “한국 전통문양의 생성형 AI 학습 데이터 구축 과정에서 나타나는 핵심적인 기술, 제도, 행위자적 요소는 무엇이며, 이들이 디지털 문화유산의 자산화에 어떠한 영향을 미치는가?”이다. 이 질문에 답하기 위해 본 논문은 사례 분석을 통해 디지털 문화재의 사회적 구성성(Social Construction of Digital Heritage)이라는 이론적 논의와 연결성을 확보하고자 한다(Tanferri, 2022). Tanferri (2022)의 연구에서 보여주듯, 디지털 문화유산은 단순한 복제가 아니라 문화적 의미의 재구성 과정이며, 이 과정에서 기술, 제도, 행위자 간의 상호작용이 새로운 문화적 정체성을 형성한다는 것을 확인할 수 있다.

아울러 메타버스, 게임, AR/VR, 디지털 아트 등 여러 응용 분야에서의 활용 가능성을 모색하고, 공공데이터 개방, 인증제도 운영, 민간 협력 모델 구축 등을 통해 AI 기반 문화데이터 생태계의 지속적 발전 방향을 제안한다. 이와 같은 접근은 생성형 AI를 활용한 문화유산의 디지털 전환과 창의적 활용을 통해 과거를 회상하고 재구성하는 새로운 방식을 제시

할 수 있을 것으로 기대된다(He et al., 2025).

연구 방법론은 George and Bennett(2005)이 제시한 탐색적 사례 연구의 체계적 분석틀을 적용하여 이론적 고찰과 정책적 함의 도출을 위한 통합적 접근을 채택하였다. George and Bennett(2005)에 따르면, 단일 사례 연구의 학술적 정당성은 사례의 독특성이 아닌 이론적 일반화 가능성에 있으며, 특히 새로운 현상에 대한 탐색적 연구에서는 사례 내 과정 추적(Process Tracing)을 통해 인과적 메커니즘을 규명하는 것이 핵심이다.

구체적인 분석 단위는 기술 요소(생성형 AI 알고리즘과 데이터 처리 기법), 제도 구조(공공데이터 거버넌스와 품질관리 체계), 행위자 역할(데이터 구축 전문가, 문화유산 전문가, AI 개발자 등)로 설정하여 이들의 상호작용을 체계적으로 분석하였다. 이러한 접근은 디지털 문화 식민주의(Digital Cultural Colonialism) 위험에 대응하는 실증적 사례를 제공하고, 문화유산 데이터 구축 과정에서 나타나는 권력관계와 의미 재구성 과정을 규명하는 데 기여한다(Simon, 2010; Kizhner et al., 2021; Tanferri, 2022).

II. 이론적 배경

2.1 한국 전통문양의 개념 및 분류체계

한국의 전통문양은 우리 민족문화의 소산으로 생활 관습, 정서, 종교, 신앙적 의미를 담고 있는 미의식

1) 본 연구 사례는 2024년 한국지능정보사회진흥원(National Information Society Agency, NIA)이 추진한 초거대 AI 확산 생태계 조성 사업의 일환으로, 국가유산진흥원과 민간기업이 협업하여 수행한 프로젝트이다(국가유산진흥원, 2024).

적 표현으로 민족의 역사를 보여준다(권기형&김성남, 2011). 전통문양은 단순한 장식 요소를 넘어 신성(神聖), 복(福), 수호, 번영 등 다양한 의미를 내포하고 있으며, 의례적 기능과 심미적 기능을 동시에 수행해 왔다(김혜진, 2019). 이러한 특성으로 인해 전통문양은 우리 민족의 집단적 가치 감정이 상징적 기호로 표현된 제2의 자연이자, 국가 문화의 시각적 언어로서 한국 문화정체성의 핵심 요소로 기능하고 있다(한국공예 · 디자인문화진흥원, 2023).

2024년 데이터 구축 사업에서는 전통문양을 형태, 용도, 시대 세 가지 기준으로 체계화하였다. 형태별

분류는 총 8종으로 구분되며, 인물문(신선, 불보살 등), 동물문(봉황, 호랑이 등), 식물문(모란, 국화 등), 자연산수문(해, 구름 등), 인공물문(불꽃, 고리 등), 문자문(수壽, 복福 등), 복합문(여러 소재 결합), 기하문(원형, 물결 등 추상 도형)으로 구성된다(국가유산진흥원, 2025).

용도별 분류는 전통문양이 적용된 실제 맥락에 따라 건축물(사찰, 서원 등), 국가유산(지정국가유산에 새겨진 문양), 생활소품(떡살, 저고리, 벼선 등)으로 구분되며, 시대별 분류는 유물의 출토 시기와 문헌기록 등을 기준으로 선사시대, 삼국시대, 고려시대,

〈표 1〉 문양 형태별 분류(국가유산진흥원, 2025)

분류	인물문	동물문	식물문	자연산수문	인공물문	문자문	복합문	기하문
정의	사람의 얼굴, 형태 또는 신선, 부처, 사천왕 등 위인의 표현 문양	인간에게 이로움과 두려움의 존재였던 동물 문양	인간에게 식량 등 풍요 제공, 자연의 아름다움을 표현한 문양	동물, 식물 제외한 해, 달, 구름 등의 자연소재 문양	동물, 식물 아닌 인공물을 형상화한 문양	특정한 글자를 연속해서 배열한 문양	여러 소재를 복합적으로 사용한 문양	선과 도형을 사용하여 표현된 문양
세부 유형	귀신문, 동자문, 불보살문, 비천문, 사람문, 신선문, 신장문, 얼굴문	가롱빈가문, 개구리문, 개형상문, 거북문, 공작문, 기러기문, 기린문, 까마귀문, 나비문, 닭문, 도철문, 돼지문, 말문, 매미문, 물고기문, 물짐승문, 박쥐문, 뱀문, 벌문, 봉황문, 사슴문, 사신문	갈대꽃문, 갯버들문, 고사리문, 국당굴문, 국화문, 꽃문, 나무문, 나무열매문, 난초문, 대나무문, 덩굴문, 마름꽃문, 매화문, 모란당굴문, 모란문, 버들문, 보상당굴문, 보상화문, 복숭화문	괴문, 구름문, 달문, 바위문, 별자리문, 빗방울문, 산수문, 산악문, 해문	고리이음문, 구슬이음문, 펜글썸문, 문살문, 불꽃문, 사슬문, 위은무늬문, 여의두문, 연꽃봉우리문, 칠보문	강령자문, 다남다자문, 덕자문, 룯자문, 만수무강문, 만자문, 범자문, 복자문, 부귀문, 수복강령자문, 수복다남자문, 수복자문, 수자문, 아자문, 회회자문	갯버들물짐승문, 구름봉황문, 구름용문, 구름학문, 꽃새문, 돌대나무문, 매화새문, 물고기파도문, 산수인물문, 소나무호랑이문, 십장생문, 연꽃당굴봉황문, 연못물짐승문, 풀벌레문	가는선문, 가는줄문, 거북등문, 굽은선문, 귀신눈문, 그물문, 격쇠문, 나란이구멍문, 나란이점문, 동그라미문, 동심원문, 마름모형문, 물결문, 물고기배문, 반동그라미문, 번개문, 빗금띠문, 빗금문, 사각형문
예시								

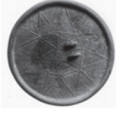
조선시대, 근현대의 다섯 시기로 구분되어 각 시대의 양식적 특징이 반영된 문양들로 분류하였다(국가유산진흥원, 2025). 이러한 분류체계는 전통문양의 다층적 의미와 문화적 맥락을 구조화하는 데 중요한

기준이 되며, 생성형 AI 학습 데이터 구축 과정에서 데이터 태깅(Tagging)과 정제(Cleaning)의 핵심 기준으로 활용되었다(국가유산진흥원, 2025).

〈표 2〉 문양 용도별 분류(국가유산진흥원, 2025)

분류	건축물	국가유산	생활소품
정의	사찰, 서원, 가옥, 석조물 등에 사용된 문양 외장재료, 내장재료로 사용된 문양 건축물에 고정되어 이동이 불가능한 객체에 사용된 문양 * 국가유산인 건축물은 건축물로 우선 분류	국가유산청이 지정하여 등록한 국가등록유산(국보, 보물, 사적, 국가민속문화유산) 시도 지정 유형문화재, 민속문화유산 국가귀속 매장문화 유물 등 * 지정된 객체에 국가유산 지정 번호 있음	인간이 생활하는데 필요로 하는 객체(생활의 편의를 위해 사용된 도구 등)에 있는 문양 과거부터 현재까지 지속적으로 사용되고 있는 객체에 있는 문양 * 국가유산 지정 번호가 없음
예시	건축물 사용(창호, 단청, 외당, 수막새기와 등) 사찰에 사용(탑, 사찰 벽 등) 왕릉, 궁궐 등(건축물 내부 설치된 석등 등)	국보 제65호 청자 기린형 뚜껑 화로(고려) 보물 제1963호 금동여래입상 청주 제82779호 칼집끝장식(청주시지정 유물)	떡살, 저고리, 버선, 치마, 배계, 허리띠, 노리개, 안경집, 별전(상평통보 등), 수귀주머니, 소반, 쫓대, 등잔, 귀걸이, 보관함, 장도, 침통 등
예시 이미지			

〈표 3〉 시대별 분류(국가유산진흥원, 2025)

시대	선사시대	삼국시대	고려시대	조선시대	근현대	시대미상
연도	BC250만년 ~ AD 37	AD 37 ~ 918년	918 ~ 1392년	1392 ~ 1897년	1897 ~ 현재	미상
주요 역사	인류 등장, 농경 시작, 독자문화권 형성	삼국의 성립과 경쟁 문화의 발전 중국문화교류 본격화	문벌 귀족의 성장 몽골 침입	성리학 도입 임진왜란 병자호란	대한제국 근대화와 도입, 일제강점기	해당 시점의 역사적 배경 미상
세부 시대	구석기, 중석기, 신석기, 청동기, 초기철기, 원삼국, 낙랑	고구려, 백제, 신라, 가야, 삼국, 통일신라, 발해, 라말여초	고려, 러말선초	조선	대한제국, 일제강점, 광복이후	-
예시						

2.2 생성형 AI 기술의 개념과 발전 동향

디지털 문화유산의 사회적 구성은 기술적 재현을 넘어 문화적 의미와 정체성의 협상 과정으로 이해되어야 한다. 특히 데이터 거버넌스는 단순 기술 관리의 범위를 넘어 문화적 권력관계와 정당성 구축의 문제로 확장되고 있는데, 디지털 문화 식민주의 현상은 특정 지역이나 문화에 대한 데이터 편향이 글로벌 AI 플랫폼을 통해 재생산되면서 문화적 다양성을 훼손하는 구조적 문제를 의미한다(Kizhner et al., 2021). 이러한 구조적 편향 문제는 생성형 AI가 문화유산 분야에 적용될 때 더욱 심화될 수 있으며, 훈련데이터에 특정 문화적 관점이 반영됨으로써 다양한 문화유산의 가치가 적절히 표현되지 못할 위험성이 존재한다(Foka and Griffin, 2024).

생성형 AI(Generative AI)는 대규모 데이터를 학습하여 텍스트, 이미지, 음성, 영상 등 새로운 콘텐츠를 자동으로 생성하는 기술로 정의되며, 최근에는 인간의 인지과 지능을 모방하여 창작의 주체로 부상하고 있다(유정재 외, 2025). 이 기술은 명령어를 통해 사용자의 의도를 스스로 이해하고 원하는 작업을 자동으로 수행하여 새로운 콘텐츠를 만들어내는 특징을 가지며, 머신러닝 기술의 정교화와 클라우드 및 소셜미디어 보편화로 인한 데이터량 증가가 비약적 발전을 이끌었다(유정재 외, 2025). 특히 텍스트-투-이미지(Text-to-Image) 분야는 문화유산 시각 자료 생성과 디지털 헤리티지 창작에 직접적인 활용 가능성을 보여주고 있으며, 최근 연구에서는 생성형 AI가 문화유산 서사 창작에서 인간과의 협업을 통해 과거를 회상하고 재구성하는 새로운 방식을 제시한다고 분석했다(He et al., 2025).

스테이블 디퓨전(Stable Diffusion)은 잠재 확산 모델(Latent Diffusion Model)을 기반으로 한 오픈소스 생성 AI로, 텍스트 프롬프트와 고해상도 이미지 간의 우수한 매핑 기능을 제공하며, 860M 파라미터 규모의 U-Net(U-shaped Convolutional Neural Network)과 CLIP(Contrastive Language-Image Pre-training) ViT-L/14(Vision Transformer Large/14) 텍스트 인코더가 결합된 구조와 VAE(Variational AutoEncoder)를 통한 잠재 공간 압축으로 계산 비용을 97%까지 절감하는 특징을 가진다(Rombach et al., 2022). 프롬프트 엔지니어링(Prompt Engineering)의 경우, 입력 프롬프트의 설계 방식에 따라 생성 이미지의 품질과 의미가 크게 달라지므로 ‘한국어(Korean)’의 문맥적 특성을 정교하게 반영하는 것이 필수적이며, 훈련 데이터에 한국 전통문양이 충분히 포함되지 않을 때 생성형 AI가 중국이나 일본 등 유사 문화의 문양을 잘못 생성하는 데이터 편향 문제도 발생할 수 있다는 점이 지적되고 있다(권구민&김현석, 2023). 이러한 편향성은 성별이나 인종에 대한 편견과 더불어 생성형 AI의 주요 안전성 이슈로 부상하면서, 불공정 요소로 작용해 사회적 불평등을 심화시키는 결과로 이어질 수 있다(유정재 외, 2025).

글로벌 시장에서는 전통문화에 특화된 독립적인 생성형 AI 모델보다는 범용 오픈소스 모델을 활용하는 추세이며, 2025년 기준 최신 대규모 언어 모델의 발전이 정체되면서 상대적으로 작은 규모의 모델에 더 많은 학습 데이터를 투입하고 전체 시스템 구조를 확장하는 방식이 차세대 모델의 주요 방향으로 제시되고 있다(유정재 외, 2025).

국내에서는 아직 전통문화에 특화된 생성형 AI 모델이 부족하며, 대부분 글로벌 오픈모델의 파인튜닝(Fine-Tuning)을 통해 활용 중이다(권구민&김현석, 2023). 국내 콘텐츠 산업에서 생성형 AI 활용을 위한 주요 해결과제로 데이터 편향성 문제, 저작권

및 윤리적 이슈, 그리고 전문 인력 부족이 지적되고 있으며, 특히 AI 전문 인력 부족 현상이 극심한 것으로 나타났다(권구민&김현석, 2023).

한국 전통문양 관련 데이터 편향 문제가 심각한 상황으로, 기존 생성형 AI 플랫폼에서 한국 전통 이미지를 학습할 데이터가 부족해 한국풍 표현이 미흡하며, 미드저니나 스테이블 디퓨전과 같은 주요 AI 모델들에서 한국 전통 이미지가 잘 구현되지 않는 이유는 서양 이미지를 중심으로 학습되었기 때문이다(국가유산진흥원, 2024). Shi et al.(2025)의 연구에서는 문화 간 메타데이터 품질 문제를 분석하며, 디지털 문화유산 관리에서 발생하는 문화적 편향과 표준화의 어려움을 지적했다.

2.3 문화데이터 구축 및 활용에 관한 선행연구

디지털 기술의 발전은 문화유산 기관의 소통 방식을 근본적으로 바꾸고 있다. Sandvik(2011)은 Web 2.0 기술과 참여형 미디어의 확산이 박물관과 문화유산 기관의 역할을 일방향적 정보 제공자에서 대화적 지식 생산자로 전환하도록 요구한다고 논의한다. 이러한 패러다임 전환에서 핵심은 관람객을 단순한 소비자가 아닌 ‘참여자(Participants)’, ‘협력자(Collaborators)’, ‘공동창작자(Co-creators)’로 인식하는 것이다(Simon, 2010).

전통문양 데이터 구축 과정도 이러한 참여적 접근이 중요한 의미를 갖는다. 문화유산 데이터 구축은 단순히 기존 문화유산을 디지털화하는 것을 넘어, AI 학습 데이터로 변환하는 과정에서 문화적 의미의

재구성과 새로운 지식 생산이 이루어지기 때문이다. 특히 이미지 캡셔닝과 메타데이터 구축 과정은 전통문양의 문화적 맥락을 현대적 언어로 번역하는 의미 재구성 과정으로 이해될 수 있다.

국내 문화데이터 구축 사례로는 문화포털과 e-뮤지엄이 대표적이며, 이들 플랫폼은 전통문양 이미지와 메타데이터를 공공누리 제1유형²⁾으로 개방하고 있다. 최근에는 인공지능 활용을 염두에 둔 메타데이터 표준화, 다국어 캡셔닝, 2D/3D 이미지 수집 등 데이터 품질과 접근성 제고를 위한 다양한 노력이 진행되고 있다. 특히 메타데이터의 다국어화 및 표준화는 문화적 맥락과 지역적 특성을 반영하는 데 핵심적인 요소로 작용하며, 고품질 데이터의 구축과 활용을 위한 중요한 기반을 형성하고 있다(Shi et al., 2025; 김나현, 2024).

이와 더불어, 한국전자통신연구원(Electronics and Telecommunications Research Institute, ETRI)과 국립중앙박물관은 지능형 헤리티지(Heritage) 플랫폼 개발 및 관련 기반 기술 연구를 활발히 수행 중이다. 구체적으로는 데이터 패브릭(Data Fabric)³⁾ 기반 아카이브, 인공지능 기반 문화유산 분석, 디지털 헤리티지 표준 수립, 생성형 AI 기반 데이터 확장 등 디지털 문화유산의 체계적 관리와 고도화된 활용을 위한 연구가 추진되고 있다(한국전자통신연구원, 2024).

해외 사례로는 유럽의 Europeana 플랫폼이 주목된다. Europeana는 유럽 각국의 디지털 문화자산을 통합하여 제공하며, 방대한 디지털 컬렉션은 물론 다국어 지원, 메타데이터의 국제 표준화, 저작권 관리, 글로벌 협력 등에서 강점을 보이고 있다(김나

2) 공공누리 제1유형은 출처만 명확히 표시하면 상업적·비상업적 목적 모두 자유롭게 활용 및 변형이 가능한 개방적인 공공저작물 라이선스이다.

3) 데이터 패브릭(Data Fabric)은 분산된 데이터를 통합·연결·관리·활용할 수 있도록 지원하는 데이터 통합 설계 개념으로, 메타데이터, 자동화, AI 기술을 활용해 데이터 접근성과 활용성을 높이는 혁신적인 방식이다.

현, 2024: European Commission, 2024). 한편, Google Arts and Culture는 AI 및 머신러닝 기술을 활용하여 세계 유물과 회화 데이터를 초고해상도로 디지털화하고, 이미지 내 세부정보 탐색, 자동 태깅, 유사 이미지 추천 등 다양한 실험적 기능을 제공함으로써 문화유산의 새로운 활용 가능성을 제시하고 있다(Kizhner et al., 2021).

전통문화 데이터의 AI 활용에 대한 학술연구도 점차 활발해지고 있다. 특히 생성형 AI 기술의 확산에 따라 데이터 편향, 저작권 이슈, 문화 정체성 왜곡 가능성 등에 대한 우려와 논의가 부상하고 있다. AI 학습 데이터의 품질과 다양성, 메타데이터의 표준화, 윤리적 고려, 문화적 편향의 내재화 문제 등은 AI 기반 문화유산 활용의 핵심 쟁점으로 자리 잡고 있다(Foka and Griffin, 2024; Tao et al., 2024). 단순한 디지털화 수준을 넘어, AI 학습에 최적화되고 고도화된 전통문화 데이터 구축은 물론, 데이터 편향의 완화와 투명성 확보, 다양한 문화적 맥락의 반영이 필수임이 학계에서 강조되고 있다(Foka and Griffin, 2024).

III. 한국 전통문양 생성형 AI 학습 데이터 구축 실증사례 분석

3.1 사례 개요 및 추진 배경

3.1.1 사례 개요

본 장에서는 2024년 초거대 AI 생태계 조성 사업의 일환으로 수행된 '2024년 데이터 구축 사업'을 실증 사례로 분석한다. 해당 사업은 국가유산진흥원과

민간기업이 협업하여 한국 전통문양을 AI 학습에 적합한 데이터셋을 체계적으로 구축하고, 문화콘텐츠 산업 및 공공서비스 분야에서의 활용 가능성을 제고하기 위해 추진되었다(국가유산진흥원, 2024).

3.1.2 추진 배경 및 필요성

한국 전통문양은 문화유산의 정체성과 상징성을 대표하는 시각자료임에도 불구하고, 기존 관련 디지털 데이터는 양적으로 부족하고, AI 학습에 적합하지 않은 형식으로 구성되어 왔다(서유덕, 2024). 특히 글로벌 AI 이미지 생성 플랫폼들인 DALL-E, Midjourney, Stable Diffusion 등에서 한국 문양 대신 중국·일본 유사 문양이 생성되는 문제는 문화 정체성 왜곡 및 디지털 주권 침해로 이어질 수 있다(He et al., 2025; Foka and Griffin, 2024).

이에 따라 전통문양을 중심으로 한 고품질 AI 학습 데이터 구축과 이에 기반한 활용 생태계 조성이 필요하다는 인식이 확산되었다(권구민&김현석, 2023).

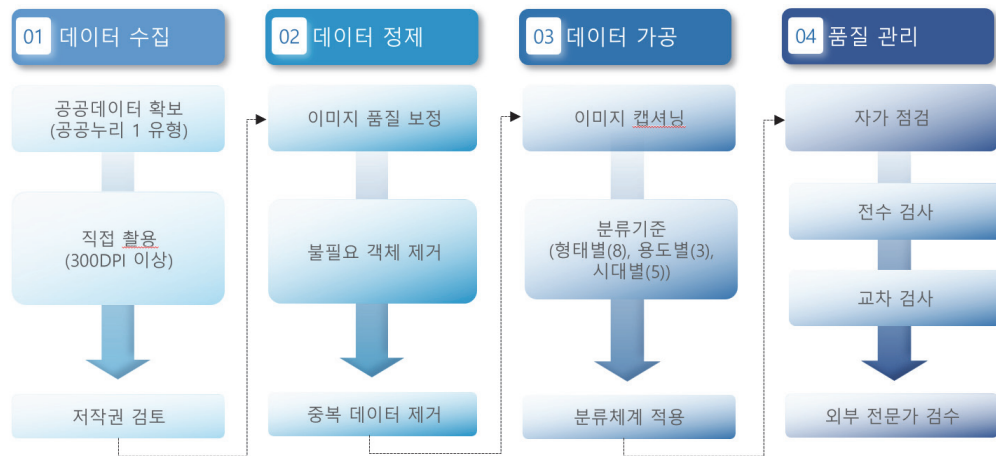
3.2 데이터 구축 프로세스

본 연구는 AI 학습용 한국 전통문양 데이터의 체계적 구축을 위해 데이터 설계·수집, 정제, 가공, 검증의 네 단계로 구성된 프로세스를 분석하였다. 각 단계는 AI 학습에 필요한 적합성을 확보하는 동시에, 전통문화의 정체성과 상징성을 유지·보존하는 것을 주요 원칙으로 하였다.

3.2.1 데이터 수집

1) 공공데이터 확보

공공데이터 확보는 국가 문화유산 기관의 개방형



〈그림 1〉 데이터 구축 프로세스(국가유산진흥원, 2025)

데이터베이스를 활용하여 수행되었다. 주요 데이터 제공 기관으로는 문화포털, e-뮤지엄, 국립중앙박물관이 포함되었으며, 이들 기관에서 공개한 전통문양 관련 이미지 데이터를 체계적으로 수집하였다. 수집 과정에서는 공공누리 제1유형 기준을 준수하여 상업적 이용과 변경이 가능한 데이터만을 선별하였다. 이는 구축된 데이터가 향후 다양한 AI 학습 목적과 상업적 활용에 제약 없이 사용될 수 있도록 하기 위한 전략적 선택이었다. 또한 각 기관이 제공하는 메타데이터(시대, 용도, 재질 등)를 함께 수집하여, 후속 데이터 정제 및 가공 단계의 효율성을 확보하였다.

2) 직접 촬영

공공데이터만으로는 충족할 수 없는 고해상도 이미지와 특수한 문양 유형에 대해서는 직접 촬영을 통해 데이터를 확보하였다. 국가유산청 협조하에 현장을 직접 방문하여 전문 장비를 활용한 고품질 촬영을 실시하였다. 촬영 대상은 생활소품(도자기, 직물, 목공예품 등), 건축물(사찰, 궁궐, 전통가옥의 문양 장식), 지정국가유산(국보, 보물급 국가유산의 문

양 요소) 등 다양한 범주로 구성되었다. 촬영 시에는 문양의 세부적 특징이 명확히 드러날 수 있도록 최소 300DPI(Dots Per Inch) 이상의 해상도를 유지하였으며, 조명과 각도를 조절하여 문양의 입체감과 색채를 정밀하게 담아내었다. 특히 현장 촬영에서는 국가유산 보존을 위한 엄격한 가이드라인을 준수하였다. 플래시 사용 금지, 최소 거리 유지, 전문 보존 담당자 입회 하에서만 촬영을 진행하는 등 문화유산의 안전성을 최우선으로 고려하였다.

3) 저작권 검토

모든 수집 데이터에 대해서는 법무 전문가와 협력하여 포괄적인 저작권 검토를 실시하였다. 이 과정에서는 저작권법, 초상권, 국가유산기본법 등 관련 법령을 종합적으로 검토하여 데이터 활용의 법적 안전성을 확보하였다. 특히 AI 학습 데이터 구축 과정에서는 여러 저작권 쟁점이 발생할 수 있다. 첫째, AI 생성물의 저작권 주체 문제로, 학습 데이터를 통해 생성된 이미지의 저작권이 원본 저작자, AI 개발자, 또는 사용자 중 누구에게 귀속되는지에 대한 명

확한 기준이 부족한 상황이다(Henderson et al., 2023). 둘째, 학습 데이터에 포함된 원본 이미지의 권리 관계 문제로, 저작권이 있는 이미지를 AI 학습용으로 사용하는 것이 공정 이용(Fair Use)에 해당하는지에 대한 논란이 지속되고 있다(Levendowski, 2018). 예를 들어, 미국의 경우 Google Books 프로젝트에서와 같이 변형적 이용(Transformative Use)으로 인정받는 사례가 있으나, 예술 작품의 경우 Andy Warhol Foundation v. Goldsmith 판결에서 보듯 상업적 목적의 변형 이용이 제한받는 경우도 있다(류시원, 2022). 본 사례에서는 이러한 법적 쟁점들을 사전에 검증하여 공공누리 제1유형 기준을 준수한 데이터만을 선별하고, AI 생성물에 대한 권리 관계를 명확히 설정함으로써 논문의 실무적 타당성과 완성도를 높이는 데 핵심적인 역할을 수행하였다.

3.2.2 데이터 정제

1) 이미지 품질 보정

저해상도, 초점 불일치, 흔들림 등 품질 저하 요인이 확인된 이미지를 제거하거나 보정하였다. 특히 저해상도 이미지는 업더닝 기반 업스케일링(Upscaling) 기법을 적용하여 해상도를 개선하였으며, 초점 불일치 및 흔들림 이미지는 디지털 영상 처리 기법을 통해 선명도를 개선하였다. 이러한 품질 보정은 AI 학습 데이터의 기술적 기준을 충족시키는 동시에, 원본 이미지가 지닌 문화적 특성을 훼손하지 않는 범위 내에서 수행되었다.

2) 불필요 객체 제거

일반적인 디지털 이미지 처리 관점에서는 전통문

양 이미지에서 배경을 단순히 시각적 잡음(Noise)으로 간주하고 제거하는 것이 보편적이다. 그러나 한국 전통문양 연구에서는 문양 자체뿐만 아니라 그것이 놓인 맥락과 배경까지 함께 고려하여 상징적·의미적 가치를 해석한다(권기형&김성남, 2011). 앞선 연구는 전통문양은 생활 관습, 종교, 정서와 결합해 형성된 상징체계이며, 배경 역시 이러한 조형성과 상징성을 강화하는 요소로 작용함을 강조한다. 예컨대 단청 문양은 직선과 곡선, 오방색의 조합뿐 아니라 건축 공간의 맥락과 결합하면서 불교적·자연적 의미를 확장한다(김혜진, 2019). 또한 전통 공예품의 문양 배경, 예를 들어 백제 금동대향로의 화염 무늬는 연화·보살상과 결합해 ‘열반’과 ‘불성’을 동시에 상징하며, 이는 배경 문양이 주체 문양과 상호작용하여 의미를 배가하는 전형적 사례라 할 수 있다(한국공예·디자인문화진흥원, 2023).

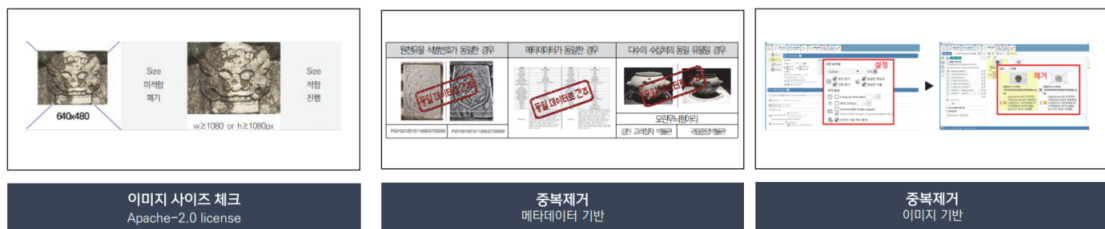
따라서 본 사례에서는 문양이 포함된 유물 전체 이미지로 배경 요소까지 담고 있는 ‘원천이미지1’⁴⁾과 해당 문양만을 분리·추출한 ‘원천이미지2’⁵⁾를 모두 보존하였다. 이를 통해 문양의 순수 형상뿐 아니라 배경이 부여하는 역사적·공간적 맥락과 상징 체계까지 분석 대상에 포함시켰다. 메타데이터에는 배경의 재질(Material), 조형 기법(Technique), 색채·질감(Color·Texture) 특성을 별도 항목으로 기록하여, 배경이 지니는 문화적 의미가 데이터 구축과 활용 전략 수립에 반영되도록 하였다. 타 객체(예: 인물, 조형물 등)는 픽셀 블러(Pixel-Level Blur) 또는 영역 마스킹(Masking) 기법으로 자연스럽게 제거되, 일부 문양이 포함된 경우에는 배경색과 동일한 색상으로 채워 문양의 식별성과 데이터 순수성을 함께 확보하였다.

4) 원천이미지1은 문양이 포함된 유물 전체 이미지며 배경 요소를 포함하고 있다.

5) 원천이미지2는 해당 문양만을 분리·추출한 이미지이다.



〈그림 2〉 원천데이터 정제 작업 예시(불필요 객체 제거)(국가유산진흥원, 2025)



〈그림 3〉 원천데이터 정제 작업 예시(중복데이터 제거)(국가유산진흥원, 2025)

3) 중복 데이터 제거

중복 데이터 제거는 메타데이터 일치 분석과 이미지 유사도 기반 비교를 병행하여 수행하였다. 메타데이터 기반 절차에서는 파일명, 촬영 일자, 유물명 등 주요 속성이 동일한 항목을 식별하여 중복 여부를 판별하였다. 이미지 기반 절차에서는 해상도, 구조적 특징, 시각적 유사성을 분석하여 동일하거나 유사한 이미지를 자동 검출한 뒤, 이 가운데 품질이 우수한 이미지만을 최종적으로 보존하였다. 이러한 절차를 통해 전체 데이터셋의 중복률을 1% 이하로 유지하였다.

3.2.3 데이터 가공

1) 이미지 캡서닝

모든 이미지는 한글과 영어로 각각 3문장 이상, 30여절 이상의 캡션을 부여하였다. 캡션은 사전에 정의한 이미지 캡서닝 규칙에 따라 작성되었으며, 문

양의 형태, 색채, 위치, 용도, 상징성, 시대적 배경 등 주요 속성을 체계적으로 기술하였다. 이러한 절차를 통해 구축된 캡션은 전통문양의 의미적 맥락을 보존하면서도 AI 학습에 최적화된 데이터로 가공되었다.

2) 분류체계 적용

전통문양 데이터는 형태별(8종), 용도별(3종), 시대별(5종) 분류기준을 전면 적용하여, 각 문양이 어떤 유형과 용도, 시대적 배경을 갖는지 명확히 구분하였다. 이러한 분류체계는 메타데이터 구조와 유기적으로 연계되어 통합적으로 정리되었으며, 각 항목은 중복 없이 독립적으로 분류되어 데이터의 체계적 관리와 검색, 활용을 용이하게 하였다.

3.2.4 품질관리

품질관리는 자가 점검, 전수 검사, 교차 검사, 외부 전문가 검수의 4단계 체계로 운영되었다. 각 단

문장	속성명	캡션 구분	캡션 내용
문장1	context_kor	객체정의 (용어)	객체의 이름과 일반화된 명칭 설명
		객체용도	건축물, 유물, 생활소품 등에서의 일반적인 활용 용도 설명
		객체기능	구조적 · 시각적 기능을 20어절 이상으로 설명
문장2	function_kor	문양추출	객체에 포함된 문양의 형태 설명
		문양위치	문양의 위치를 이미지 기준으로 설명(예시 포함)
		문양기능	문양의 구조적 · 기능적 역할을 20어절 이상으로 설명
문장3	description_kor	문양정의	문양의 정의 또는 의미를 서술
		문양묘사	선, 개수, 배치, 방향 등 요소 포함하여 300자 이상 묘사
		문양의도	문양 제작의 의도 또는 메시지 설명
문장4	symbolism_kor	색채/재질	문양에 사용된 색상과 재질을 설명(재질은 필수 입력)
		문양상징	문양이 상징하는 의미를 200자 이상으로 서술
문장5	normal_kor	통합	위 문장1~4의 내용을 자연스럽게 통합 서술
문장6	keyword_kor	키워드	문양 정보와 관련된 필수 키워드 2개 입력(예: pattern_type 등)

원천용어명	백자청화 복문대접
모양형태	문자문 원천이미지1 (객체)
문화용도	문화재 
가공매터	
1	Description 불가 치한 대접이다. 청화 청선으로 그려진 인물은 사해로 표현되었다. 물결은 양쪽으로 균형 잡힌 형태로, 물결이 바깥으로 일렁이고 있다. 배경을 흰색 도자기 표면으로, 물결의 아래 물결이 한 번 더 반복되어 전체적으로 고름처럼 보이며, 물결이 한 번의 살갗처럼 되어 있다.
2	Symblom 인물이 복자는 한국 전통에서 행복과 번영을 상징하는 중요한 문자로, 이 대접에 사용된 복문은 소유자가 복이 깃들기를 기원하는 의미를 담고 있다.
3	Keyword 문자문, 백자청화복문대접, 고려, 한국전통예술
4	Content 백자청화복문대접이라 불렀다. 그러나 원래는 백자청화복문대접이라는 말이다. 백자 그릇에 청화 색채를 칠한 문화를 복자 같은 것을 말한다. 백자 그릇 줄기에 있는 복문은 줄기로 네 방향으로 복자가 그려져 있다.
5	Funtion 백자 대접은 주로 음식을 담는 그릇으로서 사용되었지만, 기능적 목적 외에도 장례식 요소를 대체할 역할을 했다. 복문은 사용자의 복을 기원하는 의미를 지닌 대접이다.
6	Normal 불 가 치한 대접이다. 청화 청선으로 그려진 인물은 사해로 표현되었다. 물결은 양쪽으로 균형 잡힌 형태로, 물결이 바깥으로 일렁이고 있다. 배경을 흰색 도자기 표면으로, 물결의 아래 물결이 한 번 더 반복되어 전체적으로 고름처럼 보이며, 물결이 한 번의 살갗처럼 되어 있다. 그러나 원래는 백자청화복문대접이라는 말이다. 백자 그릇에 청화 색채를 칠한 문화를 복자 같은 것을 말한다. 백자 그릇 줄기에 있는 복문은 줄기로 네 방향으로 복자가 그려져 있다. 이 그릇은 주로 음식을 담는 그릇으로서 사용되었지만, 기능적 목적으로도 활용되곤 하였다.
1	Description It is a pattern featuring the character 'bok' (meaning "blessing"). The character is drawn in deep blue ink and is expressed in a traditional calligraphic style. The writing is clear and balanced, placed in the center to emphasize its importance. The background is the white surface of porcelan, which makes the blue character stand out even more. The overall design feels simple yet antique, carrying the symbolic meaning of the character 'bok'.
2	Symblom The character 'bok' is an important symbol in Korean tradition, representing happiness and prosperity. The 'bok' pattern used on this bowl carries the meaning of wishing for blessings for the owner.
3	Keyword Character pattern / Baekja Cheongwa Bokmun Bowl / 'doh', Korean traditional pattern
4	Contnt The Baekja Cheongwa Bokmun Bowl features the 'bok' character pattern. The Baekja Cheongwa Bokmun Bowl refers to a white porcelain dish with the 'bok' character drawn in blue pigment. The 'bok' character is drawn in four directions, centered on the middle of the white porcelain bowl.
5	Funtion The white porcelain bowl was primarily used for holding food, but beyond its functional purpose, it also served as an important decorative item. The 'bok' pattern added the meaning of wishing for blessings for the user.
6	Normal The character 'bok' is drawn in deep blue ink with a traditional calligraphic style, emphasized in the center of the white porcelain. The blue writing stands out against the white background, giving the piece a simple yet antique atmosphere. 'bok' symbolizes happiness and prosperity in Korean tradition, and the white porcelain carries the wish for blessings for the owner. The Baekja Cheongwa Bokmun Bowl is a white porcelain dish with the character 'bok' drawn in blue pigment, with the character positioned in four directions, centered on the bowl. This dish was mainly used for holding food but also carried significant decorative importance.

매터데이터

속성명	내용
file_name	J5_MH_1_0001
pattern_file_name	J5_MH_2_0001
format	JPG
size	1080
width	810
height	직방형
date	2024.08.23
photo_date	0519카메라
color_name	백자청화 복문대접
color_color_name	대접
material	도
usage	조각
pattern_usage	MH
pattern_usage_detail	문화재
pattern_type	M
pattern_type_desc	문자문
pattern_type_detail	복자문
pattern_symbol	기쁨과 번영과 풍요로움을 상징
year	1600
material	도자기
color	박제, 청화
category	국립중앙박물관
description	이 인물은 조선시대의 백자 청화 복문 대접을 보여준다. 이 대접의 특징은 청화 청선으로 그려진 인물이 사해로 표현되었다. 물결은 양쪽으로 균형 잡힌 형태로, 물결이 바깥으로 일렁이고 있다. 배경을 흰색 도자기 표면으로, 물결의 아래 물결이 한 번 더 반복되어 전체적으로 고름처럼 보이며, 물결이 한 번의 살갗처럼 되어 있다. 그러나 원래는 백자청화복문대접이라는 말이다. 백자 그릇에 청화 색채를 칠한 문화를 복자 같은 것을 말한다. 백자 그릇 줄기에 있는 복문은 줄기로 네 방향으로 복자가 그려져 있다. 이 그릇은 주로 음식을 담는 그릇으로서 사용되었지만, 기능적 목적으로도 활용되곤 하였다.

기준에 따라 종합적으로 점검하였다. 오류가 발견될 경우 즉시 전면 수정 또는 재작성하는 방식으로 데이터의 신뢰성과 완성도를 높였다. 특히, 데이터 수

집 단계에서는 공공데이터 활용과 현장 촬영, 저작권 검토를 통해 법적 안전성을 확보하였고, 정제 단계에서는 불필요 객체 제거 및 품질 보정으로 원본 이미지를 최적화하였다. 가공 단계에서는 한·영 캡셔닝과 용어사전 적용을 통해 AI 학습 적합성을 높였으며, 검수 단계에서는 외부 전문가가 참여하는 4단계 검수 프로세스를 통해 최종 품질을 보장하였다.

3.3 데이터 구축 과정에서 행위자 역할과 상호작용 분석

본 사업은 민간 전문기업이 사업을 주관하고 공공기관이 핵심 파트너로 참여한 사례로 민간의 주도 하에 사업이 추진되었지만, 국가유산진흥원의 참여는 프로젝트의 공공적 가치와 신뢰를 확보하고 문화유산의 정체성을 지키는 중요한 제도적 기반이 되었다. 이러한 협력 구조를 심층적으로 분석하기 위해 데이터 구축의 4대 단계와 그에 따른 세부 활동을 중심으로, 각 행위자가 어떤 역할을 수행하며 상호작용했

는지 RACI(Responsible, Accountable, Consulted, Informed) 상세 매트릭스를 통해 심층적으로 분석하고자 한다(Bernardo et al., 2024). RACI는 특정 과업에 대해 실무 책임자(Responsible, R), 최종 책임자(Accountable, A), 협의 대상(Consulted, C), 결과 공유 대상(Informed, I)을 명확히 정의하는 역할 분석 프레임워크이다. 이는 본 연구에서 제시한 “기술, 제도, 행위자 간의 복합적 상호작용”이 실제 과정에서 어떻게 구현되었는지를 보여주는 데 목적이 있다(Tanferri, 2022).

첫째, ‘데이터 수집’ 단계에서는 역할 분담에 기반한 민관협력이 이루어졌다. ‘직접 촬영’ 시 주관 사업자가 기술적 실행을 담당(R)하고, 공공기관인 국가유산진흥원은 공신력을 바탕으로 타 기관 협조 공문을 발송하는 등 실무적 지원(C, R)을 병행했다. 또한 ‘저작권 검토’에서는 주관 사업자가 최종 책임(R, A)을 지는 가운데, 공공기관인 국가유산진흥원이 법무 전문가와의 자문 과정에 함께 참여(C)하며 법적 안정성을 검토하였다.

〈표 5〉 데이터 구축 세부 프로세스별 행위자 역할 분담 (RACI) 상세 매트릭스

대분류	세부 활동	민간 전문기업 (주관사업자)	공공기관 (국가유산진흥원)	문화유산 전문가	법무 전문가	AI 개발자
01. 수집	공공데이터 확보	R, A	C	C	I	I
	직접 촬영 및 기관 협의	R	C, R	C	I	I
	저작권 검토	R, A	C	I	C	I
02. 정제	이미지 품질 보정	R, A	I	C	I	I
	불필요 객체 제거	R, A	I	C	I	I
	중복 데이터 제거	R, A	I	I	I	I
03. 가공	이미지 캡셔닝	R, A	C	C	I	I
	분류 기준 수립	R	C	A	I	I
	분류체계 적용	R, A	I	C	I	C
04. 품질	자가 점검	R, A	I	I	I	I
	전수/교차 검사	R, A	I	I	I	I
	외부 전문가 검수	R (지원)	A	R (수행)	I	R(성능검증)

둘째, ‘데이터 정제 및 가공’ 단계에서는 활동의 성격에 따라 역할을 분담했다. ‘중복 데이터 제거’와 같이 기술적 판단이 요구되는 활동은 주관 사업자가 전담(R, A)하였다. 반면, ‘불필요객체 제거’나 ‘이미지 캡셔닝’처럼 문화적 맥락에 대한 이해가 필수적인 활동에서는 문화유산 전문가의 자문(C)을 거쳤다. 특히 ‘분류 기준 수립’에서는 문화유산 전문가 그룹이 최종 결정권(A)을 가짐으로써 데이터의 학술적 체계를 도모하였다.

셋째, ‘품질 관리’ 단계는 내외부 검수를 통해 데이터의 신뢰도를 확인하는 절차로 진행되었다. 주관 사업자가 주도하는 내부 검수(R, A) 이후, 최종 ‘외부 전문가 검수’에서는 국가유산진흥원이 최종 책임(A)을 맡고 문화유산 전문가가 실무(R)를 수행하였다. 이 과정을 통해 민간에서 구축된 결과물이 공공적 자산으로 활용될 수 있는 기반이 마련되었다. 또한 이 단계에서는 AI 개발자가 CLIP/FID Score 측정 등 기술적 성능 검증을 수행(R)하여, 구축된 데이터의 AI 모델 학습 효과에 대한 실증 자료를 산출하는 역할을 수행하였다.

요약하자면, 본 사업의 추진 과정은 각 세부 활동에서 민간, 공공, 전문가의 역할이 상호 연계된 거버넌스 구조를 특징으로 한다. 이는 기술, 제도, 행위자 간의 상호작용이 디지털 문화유산의 자산화로 이어지는 과정을 구체적으로 보여주는 사례로 해석할 수 있다(George and Bennett, 2005).

3.4 데이터 구축 결과 분석

본 절에서는 2024년 데이터 구축 사업을 통해 구축된 한국 전통문양 AI 학습 데이터의 주요 결과를 제시하고, 각 평가 항목별 목표치 달성 여부 및 그 의미를 분석한다. 여기서 제시되는 다양한 평가 목표치들은 한국지능정보사회진흥원(NIA)이 추진한 『초거대 AI 확산 생태계 조성 사업』의 ‘2024년 데이터 구축 사업’ 제안요청서(Request for Proposal, RFP)에 명시된 기준에 근거하여 설정되었다. 이러한 목표치들은 RFP에 제시된 정량·정성 기준(예: 데이터 규모, 품질 검증 절차, AI 학습 활용성 등)에 기반해 산출되었으며, 전통문양 데이터의 특수성을 반영해 양적 지표(세트 수, 캡션 수), 질적 지표(중복률, 의미·구문 정확성), 활용성 지표(CLIP/FID 기준)로 구체화되었다.

모든 지표는 측정 가능성과 검증 가능성을 원칙으로 구조화되었으며, 한국지능정보사회진흥원이 주관하여 민간 수행사, 문화유산 전문가, 품질 검증 전문가 간 협의를 거쳐 확정되었다. 본 사업은 기술 구축을 넘어 공공 데이터 품질 관리 체계와 참여 행위자(데이터 구축 전문가, 문화유산 전문가, AI 개발자) 간 협업 구조라는 제도적 기반 위에서 추진되었다.

총 구축된 데이터는 24,536세트로 (원천이미지1, 가공데이터1,⁶⁾ 원천이미지2, 가공데이터2,⁷⁾ 메타데이터⁸⁾)로 이루어진 1:1:1:1 대응 구조이며, 이

6) 가공데이터1은 원천이미지1을 기반으로 하여, 객관적 묘사(Description), 상징성(Symbolism), 핵심어(Keyword), 문맥(Context), 기능(Function), 문화적 맥락(Normal)의 총 6개 항목으로 구성된 한글 서술형 정보이다. 학습, 디지털 콘텐츠 제작, 문화유산 정보 서비스 등에 활용이 가능하도록 설계되었다.

7) 가공데이터2는 가공데이터1의 6개 항목과 동일한 구조로 영문화한 데이터이다.

8) 메타데이터는 총 23개 항목으로 구성되어 있다. 주요 항목은 원천이미지1 파일명(Object File Name), 원천이미지2 파일명(Pattern File Name), 파일 형식(Format), 이미지 가로 해상도(Image Width), 이미지 세로 해상도(Image Height), 수집 방식(Collect), 촬영 일자(Photo Date), 촬영 장비(Photo Equipment), 유물 명칭(Relic Name), 유물 통칭(Relic Common Name), 시대(Era), 시대 상세(Era Detail), 문양 용도(Pattern Usage), 문양 용도 상세(Pattern Usage Detail), 문양 유형 ID(Pattern Type ID), 문양 유형(Pattern Type), 문양 유형 상세(Pattern Type Detail), 문양의 상징 의미(Pattern Symbol), 제작 연도(Production Year), 재질(Material), 색상(Color), 출처(Sources), 설명 문장(Description)이다.

와 같은 데이터 구조는 각 구성요소 간의 연계성을 명확히 하여 AI 학습, 디지털 콘텐츠 제작, 문화유산 정보 서비스 등에 활용이 가능하도록 설계되었다. 본 사업의 전체 데이터 구축량 목표는 '2024년 데이터 구축 사업' 제안요청서(RFP)에 명시된 22,000 세트였으며, 실제 구축 결과는 111.5%로 목표치를 초과하였다. 본 절에서는 구축 데이터의 형태별, 용도별, 시대별 분포와 함께 메타데이터 구성 현황을 분석하였다.

3.4.1 문양 형태별 분포

식물문 데이터가 전체의 46.8%로 가장 큰 비중을 차지하였다. 이어서 복합문(21.7%), 기하문(11.7%), 동물문(7.4%) 등의 순으로 나타났다. 구축된 전통

문양 데이터의 문양 중첩률은 56.8%를 기록하였다. 이 수치는 문양 내에 여러 종류의 다른 문양이 복합적으로 포함된 비율을 의미하며, '2024년 데이터 구축 사업'의 품질 관리 기준(50% 이상)을 상회하여 데이터 다양성이 확보되었음을 확인할 수 있다.

3.4.2 문양 용도별 분포

문양의 용도별 분포를 보면 생활소품 관련 문양이 50.8%로 가장 높은 비중을 차지하였으며, 건축물 관련 문양이 48.6%, 국가유산(국보, 보물 등)에 포함된 문양이 0.6%로 나타났다. 이러한 결과는 데이터셋이 일상생활과 밀접한 문양을 중심으로 구성되어 있음을 보여주며, 나아가 교육, 디자인, 공공서비스 분야에서의 활용 가능성을 제시한다.

〈표 6〉 형태별 전통문양 데이터 구축 현황(국가유산진흥원, 2025)

문양 형태	수량	비율(%)
식물문	11,490	46.8%
복합문	5,335	21.7%
기하문	2,868	11.7%
동물문	1,809	7.4%
인공물문	1,161	4.7%
인물문	837	3.4%
문자문	557	2.3%
자연산수문	479	2.0%
합계	24,536	100

〈표 7〉 용도별 전통문양 데이터 구축 현황(국가유산진흥원, 2025)

구분	수량	분포(%)
건축물	11,931	48.6
국가유산	157	0.6
생활소품	12,448	50.8
합계	24,536	100

〈표 8〉 시대별 전통문양 데이터 구축 현황(국가유산진흥원, 2025)

구분	수량	분포(%)
조선	7,833	31.9
삼국	7,407	30.2
근현대	5,353	21.8
고려	2,963	12.1
시대미상	934	3.8
선사	46	0.2
합계	24,536	100

3.4.3 시대별 분포

문양의 시대별 분포를 살펴보면, 조선시대가 31.9%로 가장 높은 비중을 차지하였으며, 이어 삼국시대(30.2%), 근현대(21.8%), 고려시대(12.1%) 순으로 나타났다. 선사시대 문양은 0.2%에 불과하나, 연대별 다양성 확보라는 관점에서 의미 있는 구성 요소로 평가될 수 있다.

3.4.4 메타데이터 다양성

총 132,000문장 이상의 한글·영문 병기 캡션과 1,320,000어절 이상의 텍스트 데이터를 확보하였다. 각 캡션은 문양의 형태, 색상, 용도, 위치, 상징성 등 주요 속성을 포함하도록 설계되었으며, 데이터셋의 중복률은 1% 미만으로 유지되어 고유성과 정확성이 확보되었다.

3.5 데이터 성능 평가 결과

본 사례에서는 데이터의 학습 적합성과 활용 가능

성을 검토하기 위하여 총 5개 항목(다양성, 이미지 캡셔닝 구축량, 의미 정확성, 구문 정확성, AI 모델 유효성)에 대한 성능 지표가 적용되었다. 이러한 지표와 목표치는 한국지능정보사회진흥원(NIA)에서 정의한 '2024년 데이터 구축 사업' 제안요청서(RFP)에 제시된 기준을 근거로 하였다. 요구된 목표치를 전반적으로 충족했으며, 일부 항목에서는 이를 상회하는 성과를 보인 것으로 나타났다.

3.5.1 다양성 요건

구성비 중첩률⁹⁾(캡션에 사용된 문장, 단어, 구조가 서로 다른 이미지 간에 얼마나 겹치는지 나타내는 수치)은 목표치인 50% 대비 56.77%로, 달성률 113.5%를 기록하였다. 문양 형태별로는 식물문(46.8%), 복합문(21.7%), 기하문(11.7%), 동물문(7.4%) 순으로 나타났다. 이러한 다양성 확보는 AI 모델이 특정 유형에 과적합되지 않고, 다양한 환경에서도 일관된 성능을 유지할 수 있도록 하는 핵심 요건으로 기능한다.

9) AI 학습 데이터 품질 기준을 충족하기 위해, 각 이미지는 한글 및 영어 캡션을 각각 3문장 이상, 30어절 이상으로 구성해야 하며, 그 구성의 중첩률을 달성해야 한다.

3.5.2 이미지 캡셔닝 구축량

이미지 캡셔닝은 컴퓨터 비전과 자연어 처리 기술을 융합한 대표적인 AI 응용 영역으로, 구축된 데이터는 시각 정보를 언어화하는 능력을 향상시키는 데 기여할 수 있다. 한글 및 영어 캡션의 총 구축량은 목표치 대비 224.9%로, 한글 문장 기준 278,803건(목표치 대비 211.2%), 영어 문장 기준 262,024건(목표치 대비 198.5%)으로 각각 집계되었으며, 어절 기준으로는 각각 206.6%, 283.1%를 달성하였다.

3.5.3 의미 정확성

의미 정확성 평가는 텍스트-이미지 간 의미 정합성, 주제 일치성, 번역 적정성을 기준으로 수행되었다. 그 결과 전체 평균은 목표치 95%를 상회하는 103.3%로 나타났다.

특히 번역 적정성은 문화유산 전문 번역업체인 월시스가 수행한 캡션 번역 결과에 대해, 한국정보통신

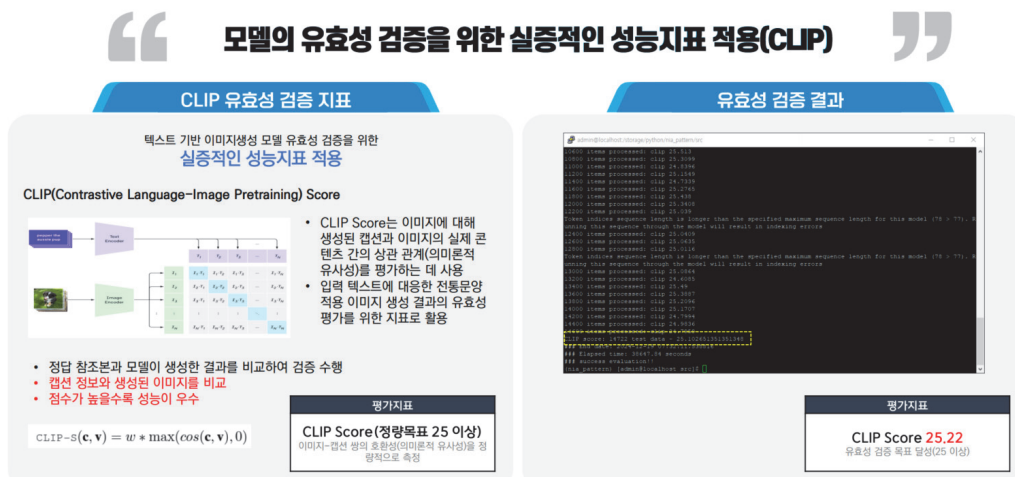
신기술협회(Telecommunications Technology Association, TTA)가 적정성 검증을 수행한 결과 98.95%의 정확도를 기록하였으며(목표 대비 104.2%), 이는 번역 품질의 신뢰성을 입증하는 지표로 해석될 수 있다.

3.5.4 구문 정확성

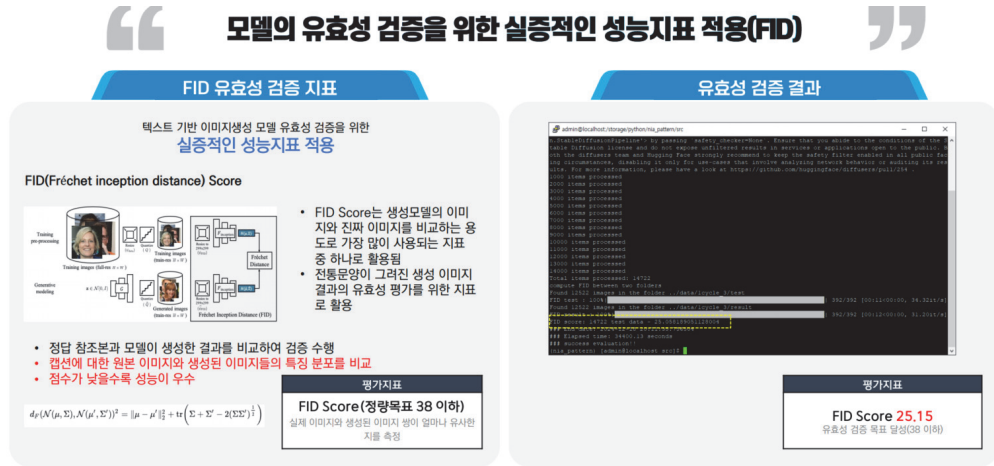
구문 정확성은 문장의 문법적 구조와 형식의 정합성을 기준으로 평가되었다. 그 결과, 목표치 99.5%를 상회하는 100.5%를 기록하였다. 이는 데이터셋 내 문장이 문법적으로 안정적이며, 자연스럽게 유창한 구성을 유지하고 있음을 의미한다.

3.5.5 AI 모델 유효성

AI 이미지 생성 성능 평가는 CLIP Score와 FID(Frechet Inception Distance) 지표를 활용하였다. CLIP Score는 이미지와 텍스트 설명 간의 의미



〈그림 5〉 모델의 유효성 검증을 위한 CLIP 검증 결과 (국가유산진흥원, 2025)



〈그림 6〉 모델의 유효성 검증을 위한 FID 검증결과 (국가유산진흥원, 2025)

적 일치도를 수치로 평가하는 지표이며, FID는 생성 이미지와 실제 이미지 간의 분포 차이를 측정하여 이미지 품질을 평가하는 대표적 척도이다. 분석 결과, CLIP Score 목표치 25 이상에 대해 25.22(목표치 대비 100.9%)를 기록하였고, FID 목표치 38 이하에 대해 25.15(목표치 대비 151.1%)를 기록하였다. 이는 본 사례에서 구축한 데이터셋이 생성 이미지의 품질과 일관성 측면에서 우수함을 보여준다.

3.5.6 종합 평가

본 사례에서 구축된 데이터는 다양성, 구조화, 정합성, AI 적합성 등 주요 지표에서 목표치를 전반적으로 상회하였다. 특히 이미지 캡셔닝 구축량(224.9%)과 AI 모델 유효성(126.0%) 항목에서 상대적으로 높은 수치를 기록하였다. 이러한 성과는 한국지능정보사회진흥원(National Information Society Agency, NIA)이 수행한 외부 평가에서도 긍정적으로 검증되었으며, 종합적으로는 “매우 우수” 등급에 해당하는 것으로 확인되었다.

3.6 실증사례의 시사점

본 실증사례 분석은 George and Bennett(2005)이 제시한 과정 추적(Process Tracing) 방법을 적용하여 데이터 구축 과정에서 나타나는 인과적 메커니즘을 체계적으로 추적하였다. 과정 추적은 “연구자가 역사적 기록, 아카이브 문서, 인터뷰 전사본 등의 자료를 검토하여 이론이 가정하거나 함의하는 인과적 과정이 실제로 해당 사례의 중간 변수들의 순서와 값에서 나타나는지를 확인하는 방법”이다(George and Bennett, 2005).

본 연구는 데이터 구축 과정의 각 단계(수집-정제-가공-검증)에서 기술적 결정, 제도적 제약, 행위자 간 상호작용이 어떻게 최종 데이터의 품질과 문화적 정체성 보존에 영향을 미쳤는지를 추적함으로써, 디지털 문화유산 자산화의 핵심 메커니즘을 규명하였다.

3.6.1 디지털 문화유산의 자산화 기반 마련

본 사례는 전통문양을 단순 보존 대상에서 디지털

기반 활용 자산으로 전환하는 토대를 마련하였으며, 이는 문화유산 관리 정책의 패러다임 전환을 의미한다. 기존의 문화유산 보존 방식이 원형 그대로의 물리적 보존에 집중했다면, 이번 사업은 디지털 변환을 통한 능동적 활용 모델을 제시했다. 특히 24,536세트의 체계적 데이터 구축은 전통문화의 정체성을 보존하면서도 생성형 AI 기술과의 융합 가능성을 실증적으로 입증하였다.

이 실증사례는 단순한 디지털화를 넘어 AI 기반의 문화유산 복원과 재창조 모델을 성공적으로 구현했다. 전통 연회, 춤, 음악을 디지털화하여 지속적으로 보존할 수 있는 기반을 마련했으며, AI 기반 아카이브와 콘텐츠 제작을 통해 젊은 세대와의 접점을 확대할 수 있는 방향을 제시했다. 또한 이 사업을 통해 제작된 복어 모양 목공예 상품 등 6종의 AI 생성 문화상품 시제품은 전통문화가 단순한 보존 대상이 아닌 경제적 가치를 창출하는 자산임을 입증했다. 이는 그동안 보존과 전승 개념의 틀을 깨고 문화적 자산이 경제적 자산이 되는 선순환 구조를 만드는 가능성을 제시했으며, 전통문화의 현대적 방식 재해석을 통한 산업화 모델의 성공적 사례로 평가된다.

3.6.2 국제적 경쟁력 확보

한·영 병기 이미지 캡처링, 표준화된 분류체계, 4단계 품질관리 체계를 통해 구축된 데이터는 글로벌 AI 학습 플랫폼과의 호환성을 확보할 수 있는 수준에 도달하였다. 이는 한국 전통문양의 국제적 확산과 활용 기반을 마련하는 데 기여한다. 학습데이터를 활용함으로써 기존 글로벌 AI 플랫폼에서 한국 문양 대신 중국이나 일본의 유사 문양이 생성되는 문제를 개선할 수 있었으며, 이는 문화적 디지털 주권 확립의 모범 사례로 평가된다. 이러한 성과는 각국이 자국

의 문화적 특성을 반영한 AI 모델 개발에 주력하는 글로벌 트렌드 속에서 한국이 선도적 위치를 확보할 수 있는 기반을 마련했으며, 문화 정체성 보호와 디지털 주권 확보를 위한 실질적인 방안을 제시한다.

3.6.3 문화콘텐츠 산업으로의 확장성

게임, 메타버스, 디지털 아트, AR/VR, NFT(Non-Fungible Token) 등 다양한 분야에서 활용 가능한 기반 데이터가 성공적으로 구축되었으며, 전통문양의 현대적 재해석 및 콘텐츠 융합 모델 개발이 가능함을 입증하였다. 구축된 데이터는 디지털 박물관, 미디어 아트, 패션 디자인, 게임 및 메타버스 콘텐츠 개발, 교육 자료 제작 등에 즉시 활용될 수 있는 수준의 품질을 확보하였다. 이 데이터는 전통 문양의 현대적 재해석뿐만 아니라, 다양한 사업 분야와의 융합을 통해 새로운 문화콘텐츠와 비즈니스 모델 창출의 기반을 제공할 것으로 기대된다. 특히 생성형 AI 모델을 활용한 전통 문양의 디지털 재생산 및 창작, NFT 등 신산업 영역에서의 활용을 통한 문화콘텐츠 산업 전반으로의 확장성을 시사함과 동시에 전통적 한계를 초월한 새로운 창작 패러다임이 형성될 가능성을 제시하였다.

3.6.4 AI 생태계 조성을 위한 정책 기반 마련

공공데이터 개방, Open API(Application Programming Interface) 연계, 전통문양 인증제도 운영 등은 AI 생태계 활성화의 실질적 기반이 될 수 있으며, 특히 2025년 6월 AI허브를 통한 데이터 공개는 정부의 공공데이터 평가 체계 개편과 맞물려 AI 친화적이고 고가치 데이터 개방의 모범 사례로 평가된다. 이러한 데이터 개방과 품질관리 체계는 AI

기반 공공서비스의 효율성과 효과성을 높였으며, 사업이 '매우 우수' 등급을 획득한 성과는 산업 생태계 조성에 있어 실질적이고 구체적인 성공 사례로 작용했다.

IV. 생성형 AI 기반 활용 방안

본 장에서는 구축된 한국 전통문양 생성형 AI 학습 데이터를 기반으로, 선행 연구 및 정책자료에서 제시된 활용 사례와 본 연구의 실증 데이터를 종합적으로 분석하여, 문화콘텐츠 산업, 공공서비스 및 교육 분야에서의 활용 방안과 데이터 확산 및 생태계 조성 전략을 제시한다. 실현 가능한 응용 시나리오를 중심으로 구체화하고 실제 적용 사례를 중심으로 분석하였으며, 각 방안의 타당성은 관련 선행 연구와 실증사례를 근거로 논의하였다.

4.1 문화콘텐츠 산업 활용 방안

한국 전통문양 데이터는 시각적 상징성과 정체성을 갖춘 고유 자산으로, 생성형 AI 기술과 결합할 경우 다양한 문화콘텐츠로의 확장 가능성을 지닌다(노형후&박상명, 2023). 특히 전통문양의 시각적 요소는 현대적 디지털 콘텐츠와 융합되어 새로운 문화적 가치를 창출할 수 있는 잠재력을 보유하고 있으며, 이러한 융합은 한국 문화의 글로벌 확산과 디지털 문화유산의 지속가능한 발전에 기여할 수 있다(Wu et al., 2024). 이러한 이론적 가능성이 현실적으로 구현된 대표 사례가 2025년 6월 20일 넷플릭스를 통해 전 세계에 공개된 미국 소니픽처스 애니메이션의 뮤지컬 판타지 애니메이션 영화 '케이팝 데몬 헌터스'이다. 이 작품은 K-팝 아이돌 그룹이 악령과 싸우는

퇴마사로 활동한다는 독특한 설정을 통해 K-팝의 화려한 퍼포먼스와 한국의 건축물, 전통문양, 민화, 무속 문화를 창의적으로 접목하였다. 공개 직후 41개국에서 넷플릭스 1위를 차지하며 전통문화와 현대 대중문화의 성공적 융합 모델을 제시했다.

또한 게임·메타버스 및 디지털 아트 융합 콘텐츠로서 전통문양은 캐릭터 의상, 배경 텍스처, 아이템 디자인 등으로 재해석되어 가상환경에 적용될 수 있으며, 이를 통해 이용자의 몰입도와 문화적 정체성을 강화할 수 있다(노형후&박상명, 2023). 실제로 2023년 메타버스 엑스포에서 공개된 제주목관아 배경 VFX(Visual Effects) 스튜디오와 전통 활쏘기 VR 게임은 전통문양을 활용한 몰입형 사례로 평가받아 콘텐츠의 상업·문화적 성공 가능성을 입증하였다(한국공예·디자인문화진흥원, 2023). 생성형 AI 기술은 이러한 제작 과정에서 자동화된 자산 생성을 통해 개발 비용 절감과 품질 향상을 동시에 달성할 수 있어, 전통문양 데이터의 활용 범위를 게임 캐릭터 의상 디자인, 건축물 장식, 가상 공간 배경 텍스처 등으로 확장하는 기반을 제공한다(유정재 외, 2025, 박이숙&고대영, 2024).

전통문양을 재구성한 AI 생성 이미지는 NFT 형태로 글로벌 유통되어 디지털 자산화가 가능하며, 이는 한국 고유 문양을 새로운 시장에 소개하는 효과적인 수단이 된다(Wu et al., 2024). 간송미술관의 훈민정음 해례본 NFT는 개당 1억 원에 100개를 발행하여 80개 이상이 판매되는 성과를 거두며 전통문화의 상업적 가치와 글로벌 확산 채널로서 가능성을 입증하였다(한국공예·디자인문화진흥원, 2023). Stable Diffusion과 같은 고해상도 이미지 합성 모델은 전통문양의 정교한 디테일과 복잡한 패턴을 정확히 재현하면서도 창의적 변형을 가능하게 해, 전통과 현대가 공존하는 디지털 아트 창작을 지원한다

(Rombach et al., 2022).

AR/VR 기술과 결합하면 전통문양을 박물관·전시관에서 물입형 체험 콘텐츠로 제공하여 관람객의 참여도와 학습 효과를 동시에 높일 수 있다(He et al., 2025). 3D 레이저 스캐닝, 깊이 이미지, 포토그래메트리 등 다양한 3D 모델링 기법을 활용하면 고품질의 입체 문화유산 데이터를 구축할 수 있으며, 이를 통해 전통문양을 단순한 2D 이미지가 아닌 상호작용 가능한 디지털 콘텐츠로 발전시킬 수 있다(Yusri et al., 2022). 이러한 기술적 기반은 문화유산 보존과 교육, 체험 콘텐츠 개발의 새로운 지평을 열어준다.

4.2 공공서비스 및 교육 활용 방안

디지털 문화유산 데이터를 공공과 교육 분야에 적용하면 문화유산의 대중화 및 학습 효율성을 획기적으로 높일 수 있으며, 이를 통해 전통문양과 같은 시각자료에 대한 접근성을 강화하고, 특히 디지털 네이티브 세대의 문화유산 이해와 참여를 촉진하는 중요한 수단이 된다(김나현, 2024).

전통문양 AI 데이터를 교육 콘텐츠로 제작할 때는 초·중·고등학교 및 평생교육 과정에서 활용할 시각자료, 시뮬레이션 기반 학습모듈, 온라인 학습 플랫폼 콘텐츠 개발이 가능하다(Shi et al., 2025). AI 기술을 활용하면 학습자의 개인별 흥미와 참여도를 높이는 맞춤형 학습 경험을 제공할 수 있으며(이한신&김관수, 2019), 전통문양을 다양한 형태로 변형·재현함으로써 학습자들이 능동적으로 문화유산을 탐구할 기회를 제공한다(Shi et al., 2025).

공공정보 서비스 및 홍보 콘텐츠 분야에서는 문화유산 데이터를 자동화 콘텐츠 생성 기술과 결합해 공공기관 웹사이트, 박물관 전광판, 모바일 애플리케이션

이전 등에 적용함으로써 폭넓은 홍보와 교육 효과를 거둘 수 있다(김나현, 2024). ETRI와 국립중앙박물관이 공동 개발한 지능형 헤리티지 플랫폼을 통해 인천국제공항 1터미널의 반가사유상 디지털 콘텐츠 및 국립중앙박물관 ‘역사의 길’ 광개토대왕릉비 디지털 전시가 성공적으로 진행되었으며(한국전자통신연구원, 2024), 이를 통해 고품질 문화유산 디지털화와 실감형 콘텐츠 제작이 가능해져 전통문양 활용 범위가 크게 확장되었다(한국전자통신연구원, 2024).

공공기관 및 디자인 기업은 텍스트 기반 명령어만으로 문화유산 디자인을 생성할 수 있는 생성형 AI 창작 도구를 개발하여 콘텐츠 제작의 효율성과 확장성을 제고할 수 있다(Foka and Griffin, 2024). 생성형 AI는 창작자의 역할을 보조·확장하며, 전통문양의 정확한 재현과 창의적 변형을 동시에 가능하게 해 문화유산 분야에서 혁신적 도구로 자리잡고 있다(유정재 외, 2025). 그러나 AI 활용 시 문화적 편향과 정확성 문제를 최소화하기 위해 한국 전통문양의 특성과 의미를 보존하는 균형잡힌 접근이 필수적이다(Foka and Griffin, 2024; 유정재 외, 2025).

4.3 데이터 확산 및 생태계 조성 전략

생성형 AI의 안정적 확산과 활용을 위해서는 고도화된 데이터베이스 구축을 중심으로, 산업 목적에 부합하는 AI 알고리즘 설계, 신뢰성 있는 학습 데이터 확보, 중소기업 지원, 전문 인력 양성, 제도적 기반 정비 등이 종합적으로 추진될 필요가 있다(권구민&김현석, 2023). 이러한 전략은 데이터의 단순한 공개를 넘어, 산업 전반에서 지속가능한 생태계를 조성하기 위한 구조적 접근으로 이해될 수 있다.

AI허브와 공공데이터포털에서 전통문양 데이터셋을 개방하고, Open API 형태로 기업과 개발자에게

제공하면 민간 분야 활용이 활성화될 수 있다(권구민 & 김현석, 2023). 이러한 공공데이터 개방은 디지털 경제 시대의 핵심 동력이 되며, 특히 문화유산 데이터는 창조경제와 문화산업 발전에 직접 기여한다(김나현, 2024). 유럽연합이 운영하는 Europeana 프로젝트는 36개국 도서관 및 미술관 2,300여 곳이 참여해 3,200만 건 이상의 디지털 문화유산 자료를 서비스하는 성공 모델을 제시했으며, 이는 한국 공공데이터 개방 정책에 중요한 시사점을 제공한다(European Commission, 2024).

한국 문화유산의 정체성과 품질을 확보하기 위해서는 단순한 기술적 정확성을 넘어 문화적 진정성과 의미 보존을 포함하는 전통문양 인증제도를 도입해야 한다(Foka and Griffin, 2024). 특히 글로벌 AI 모델이 서구 중심 편향을 보이는 상황에서, 한국 문화유산의 고유 특성과 문화적 맥락을 정확히 반영하는 인증 체계는 문화적 디지털 주권 확보의 핵심 수단이다(Tao et al., 2024). 또한 지역 문화단체, 공예작가, 디자인 스타트업과의 협업을 통해 전통문양 기반 공동 프로젝트 추진을 통해 지역문화 산업 활성화에 기여할 수 있다. 한국공예·디자인문화진흥원이 2023년 전통문양 산업활용 기반구축 사업에서 ITKUNA, MWM, Pisca 등 민간 브랜드와 협업하여 타월, 플레이트, 쿠션, 인형 등 생활용품을 개발하고, 공항철도 승강장 등 공공공간에 전시·판매한 사례는 지역 맞춤형 지속가능 모델을 제시한다(한국공예·디자인문화진흥원, 2023). 이러한 접근은 전통문양 데이터를 기술 자원으로 활용하는 것을 넘어 지역 공동체의 문화적 정체성 강화와 경제적 가치 창출을 동시에 달성하는 방안을 보여준다.

V. 결론 및 시사점

한류는 전 세계 119개국에서 2억 2,500만 명의 팬덤을 형성하며 글로벌 대중문화로 자리매김하고 있다. K-팝과 드라마를 시작으로 한식, 뷰티, 패션, 게임 등 생활문화 전반으로 확장된 한류는 외국인의 한국 방문과 전통문화 체험 욕구를 함께 견인하고 있다(한국국제교류재단, 2024). 특히 2024년 상반기 국립중앙박물관 외국인 관람객 수는 94,951명으로 역대 최고치를 기록하며 전년 대비 34.5% 증가했다. 이처럼 한류에 힘입은 국립중앙박물관 등 문화유산 기관에 대한 관심 증가는 한국 전통문양의 디지털 자산화 및 AI 학습 데이터 구축의 당위성을 더욱 높이고 있다.

본 연구는 초거대 인공지능(Hyper-scale AI) 시대를 대비하여 한국 전통문양의 생성형 AI 학습 데이터 구축 사례를 실증적으로 분석하고, 이를 기반으로 한 문화콘텐츠 산업 활용 방안과 데이터 생태계 조성 전략을 종합적으로 제안하였다. 2024년 국가유산진흥원이 수행한 「한국 전통문양 생성형 AI 학습 데이터 구축 사업」의 결과를 기반으로 데이터 수집, 정제, 가공, 품질관리 등 전 과정을 실증적으로 분석하였으며, 총 24,536세트의 데이터를 형태별, 용도별, 시대별로 분류하고 한글 및 영어 병기의 이미지 캡션을 부여하여 AI 학습 적합성을 확보하였다. 이를 바탕으로 게임, 메타버스, 디지털 아트, AR/VR 등 문화콘텐츠 산업 분야뿐만 아니라 교육, 공공서비스 영역에서의 활용 방안과 공공데이터 개방, 인증제도 운영, 민간 협력모델 구축 등 데이터 생태계 확산 전략을 종합적으로 제시하였다.

본 연구는 단일 사례를 대상으로 이루어졌으나, AI 시대에 전통문화 데이터를 구축하고 활용하는 과정

에서 나타나는 데이터 거버넌스와 문화 정체성 구성 등 사회과학적 쟁점을 심층적으로 탐구함으로써 학술적으로 기여한다. 특히, 한국이 문화적 디지털 주권을 확보하고, 전통문화의 현대적 계승을 위한 성공적인 정책 모델을 제시했다는 점에서 전통문양 외 다른 전통문화 요소로의 일반화 가능성을 지닌다.

연구의 이론적 기여는 Tanferri(2022)가 제시한 ‘디지털화 과정에서의 사회적 구성성’ 관점을 생성형 AI 시대의 맥락으로 확장한 데 있다. 특히, AI 학습 데이터 구축 과정에서 드러나는 기술-제도-행위자 간 복합적 상호작용이 전통문화의 현대적 재해석을 가능하게 하고, 이를 통해 디지털 문화유산이 새로운 문화적 자산으로 자리매김하며 글로벌 확산의 잠재력을 형성하는 과정을 실증적으로 규명하였다.

또한 George and Bennett(2005)의 사례 연구 방법론을 적용하여, 이론적 일반화가 가능한 분석 틀을 제시하였다. 본 연구에서 규명된 데이터 거버넌스와 문화 정당성 구축 메커니즘은 한국의 전통문화 뿐만 아니라 다른 문화권의 전통문화 디지털화 사업에도 적용 가능한 이론적 통찰을 제공한다.

본 연구의 실증 분석은 단일 사례 연구로서 데이터 구축의 전 과정을 실질적으로 조망하고 성과를 검토하였으나, 데이터 구축 과정과 구축물의 질적·양적 성과에 집중되어 실제 서비스 개발, 산업적 적용, 사용자 수용도 분석 등 2차적 활용 효과에 대한 실증적 평가가 미흡하였다는 한계가 존재한다. 또한 전통문양이라는 특정 주제를 중심으로 한 사례 연구로서 다른 유형의 전통문화 데이터나 타국 문화유산 데이터와의 비교 일반화에는 제한이 있으며, 향후 연구에서는 이러한 한계를 보완할 수 있는 다각적 접근이 필요하다.

향후 연구 과제로는 한국 전통문양 학습 데이터를 기반으로 한 이미지 생성 전·후 결과의 정량적 비교

와 이미지의 문화 정체성 반영 정도, CLIP Score, FID 등 품질 평가 지표를 활용한 사용자 인식 분석을 포함한 실험 설계 연구가 필요하다. 구축된 데이터를 활용한 게임, 메타버스, NFT, 교육 콘텐츠 등의 실제 개발 사례를 대상으로 사용자 반응, 산업적 가치 창출, 수용도, 지속성 등을 검토하는 사례 기반 활용 연구가 요구되며, 현재 2D 이미지 중심의 데이터에서 나아가 3D 스캔 및 모델링 기술을 활용한 입체 전통문양 구축과 이를 기반으로 한 공간 기반 AI 학습 가능성에 대한 탐색이 중요하다. Europeana, Google Arts and Culture 등 해외 문화유산 데이터베이스와의 표준 호환성 분석 및 공동 AI 모델 학습 실험을 통해 한국 전통문양 데이터의 글로벌 활용성 확대와 국제 공동연구 기반 확보가 가능할 것으로 전망되며, 이는 생성형 AI 시대에 대응한 전통문화 데이터의 디지털 자산화 방향성과 실증 사례를 제시하는 본 연구의 성과를 바탕으로, 한류의 글로벌 확산이라는 시대적 기회를 활용하여 전통문화의 현대적 계승과 발전을 도모할 수 있는 중요한 기반을 마련했다는 점에서 그 의미가 크다. 후속 연구를 통해 활용성 검증과 산업 연계 성과 분석이 지속적으로 확장될 수 있을 것으로 기대된다.

참고문헌

- 국가유산진흥원. “AI의 우리나라 전통문양 생성 돕는다.” 국가유산진흥원 보도자료. <https://www.kh.or.kr/short/F01h>, 2024년 10월 접속.
- (Korea Heritage Agency. “AI helps generate traditional korean patterns,” Korea Heritage Agency Press Release, <https://www.kh.or.kr/short/F01h>, retrieved October 2024.)

- 국가유산진흥원 (2025). "2024년 한국 전통문양 생성형 AI 학습 데이터 구축 사업 결과보고서," **국가유산진흥원 보고서**.
- (Korea Heritage Agency (2025). "2024 Project report on the construction of generative AI training data on korean traditional patterns," *Korea Heritage Agency Report*.)
- 국립중앙박물관. "국립중앙박물관 상반기 외국인 관람객 9만 명, 역대 최고 돌파," 국립중앙박물관 보도자료, <https://www.museum.go.kr/MUSEUM/contents/M0701040000.do?schM=view&catCustomType=post&catId=93&arcId=21230&cp=27&sv=>, 2024년 10월 접속.
- (National Museum of Korea. "Record-breaking 90,000 foreign visitors at the National Museum of Korea in the first half of 2024," National Museum of Korea Press Release, <https://www.museum.go.kr/MUSEUM/contents/M0701040000.do?schM=view&catCustomType=post&catId=93&arcId=21230&cp=27&sv=>, retrieved October 2024.)
- 권구민, 김현석 (2023). "콘텐츠산업의 생성형 AI 활용 이슈와 대응 과제," **KOCCA FOCUS**, 150.
- (Kwon, G. M. and Kim, H. S. (2023). "Issues and challenges in utilizing generative AI in the content industry," *KOCCA FOCUS*, No.150.)
- 권기형, 김성남 (2011). "한국 전통문양을 응용한 업스타일 헤어작품 연구," **한국디자인문화학회**, pp.26-40.
- (Kwon, K. H. and Kim, S. N. (2011). "A study on up-style hair works using korean traditional patterns," *Journal of the Korean Society of Design Culture*, pp.26-40.)
- 김나현 (2024). "한국 디지털 문화유산(Digital Heritage) 관리 방안: 유럽 유로피아나(Europeana)와 한국 e뮤지엄 사례를 중심으로," **디지털콘텐츠학회논문지**, 제25권 8호, pp.2319-2330.
- (Kim, N. H. (2024). "Management strategies for digital heritage in korea: A case study of korea's eMuseum and Europe's Europeana," *Journal of Digital Contents Society*, 25(8), pp.2319-2330.)
- 김혜진 (2019). "한국 전통 단청 문양을 활용한 패션디자인 연구," **한국의류학회지**, 제43권 2호, pp.215-228.
- (Kim, H. J. (2019). "A study on fashion design using korean traditional dancheong patterns," *Journal of the Korean Society of Clothing and Textiles*, 43(2), pp.215-228.)
- 노형후, 박상명 (2023). "문화·관광 분야에서의 생성형 AI 지자체 활용 방안," **한국지역정보개발원**, 제23권 6호.
- (Noh, H. H. and Park, S. M. (2023). "Strategies for local government use of Generative AI in the culture and tourism sectors," *Korea Regional Information Development Institute*, 23(6).)
- 박이숙, 고대영 (2024). "국내 고성장 ICT 서비스기업의 효율성에 관한 연구," **경영학연구**, 제53권 제2호, pp.249-263.
- (Park, Y. S. and Ko, D. Y. (2024). "A Study on the Efficiency of High-Growth ICT Service Companies," *Korean Management Review*, 53(2), pp.249-263.)
- 류시원 (2022). "공정이용 판단의 고려요소로서 '변형적 이용'의 한계," **계간 저작권**, 140, pp.5-50. DOI:10.30582/kdps.2022.35.4.5
- (Ryu, S. W. (2022). "The limitations of transformative use as a consideration in fair use judgment," *Quarterly Copyright*, 140, pp.5-50.)
- 서유덕. "문화유산에 AI 접목...디지털 전환 앞장," 정보통신신문, <https://www.koit.co.kr/news/articleView.html?idxno=122817>, 2024년 10월 접속.
- (Seo, Y. D. "Integrating AI into cultural heritage... Leading digital transformation," *ICT News*, <https://www.koit.co.kr/news/articleView>.

- html?idxno=122817, retrieved October 2024.)
에이치씨아이플러스. “한국 전통 문양 데이터-인공지능 학습 데이터 교육 영상,” NIA AI허브, <https://www.youtube.com/watch?v=GPRAwmHo0yY>, 2025년 5월 접속.
- (HCI Plus. “Educational video on korean traditional pattern data for AI training,” NIA AIHub, <https://www.youtube.com/watch?v=GPRAwmHo0yY>, retrieved May 2025.)
- 이한신, 김판수 (2019). “소비자의 기술수용과 저항이 인공지능(AI) 사용의도에 미치는 영향,” **경영학연구**, 제48권 2호, pp.423-447.
- (Lee, H. S. and Kim, P. S. (2019). “The effect of consumer’s technology acceptance and resistance on intention to use of artificial intelligence(AI),” *Korea Management Review*, 48(2), pp.423-447.)
- 정명희 (2023). “한류4.0과 전라북도 전통문화자원 연계 필요,” **전라북도발전연구원 이슈 브리핑**, 제78권
- (Jeong, M. H. (2023). “The need to link hallyu(the korean wave) 4.0 with traditional cultural resources of jeollabuk-do,” *J-Think Issue Briefing*, 78.)
- 정보통신정책연구원 (2025). “생성형 AI가 미디어 분야에 미칠 영향,” **KISDI Premium Report**.
- (Korea Information Society Development Institute (2025). “The impact of generative AI on the media sector,” *KISDI Premium Report*.)
- 한국국제교류재단 (2024). “2023 지구촌 한류현황,” **한국국제교류재단**.
- (Korea Foundation (2024). “2023 Global hallyu(the korean wave) Status,” *Korea Foundation*.)
- 한국공예·디자인문화진흥원 (2023). “2023 전통문양 산업활용 기반구축 결과보고서,” **한국공예·디자인문화진흥원**.
- (Korea Craft & Design Foundation (2023). “2023 Result report on the establishment of industrial utilization infrastructure for traditional patterns,” *Korea Craft & Design Foundation*.)
- 한국전자통신연구원. “ETRI, 문화유산에 디지털로 혼 불어 넣는다,” ETRI Webzine, 242, <https://www.etri.re.kr/webzine/202408/sub05.html>, 2024년 10월 접속.
- (ETRI. “ETRI breathes digital life into cultural heritage,” ETRI Webzine, 242, <https://www.etri.re.kr/webzine/202408/sub05.html>, retrieved October 2024.)
- 허명숙, 천면중 (2021). “디지털 성숙도 진단모형 개발과 적용을 통한 디지털 트랜스포메이션 준비성에 관한 연구: 석유화학산업 S회사의 사례를 중심으로,” **경영학연구**, 제50권 제1호, pp.81-114.
- (Heo, M. S. and Cheon, M. J. (2021). “A Study on the Digital Transformation Readiness Through Developing and Applying Digital Maturity Diagnosis Model: Focused on the Case of a S Company in Oil and Chemical Industry,” *Korean Management Review*, 50(1), pp.81-114.)
- Bernardo, B. M. V., Mamede, H. S., Barroso, J. M. P. and dos Santos, V. M. P. D. (2024). “Data governance & quality management – Innovation and breakthroughs across different fields,” *Journal of Innovation & Knowledge*, 9, 100598. <https://doi.org/10.1016/j.jik.2024.100598>
- European Commission. Summary report of the public consultation on Europeana, Europe’s digital platform for cultural heritage, <https://digital-strategy.ec.europa.eu/en/library/summary-report-public-consultation-europeana-europes-digital-platform-cultural-heritage>, retrieved October 2024.
- Foka, A. and Griffin, A. (2024). “AI, cultural heritage, and bias: Some key queries that arise from

- the use of GenAI," *Heritage*, 7(11), pp.6125-6136.
- George, A. L. and Bennett, A. (2005), *Case studies and theory development in the social sciences*, Cambridge, MA, MIT Press.
- He, Z., Su, J., Chen, L., Wang, T. and Ray, L. C. (2025). "I recall the past": Exploring how people collaborate with generative AI to create cultural heritage narratives, *arXiv preprint*, arXiv:2501.00359.
- Henderson, P. Li. X., Jurafsky, D., Hashimoto, T., Lemley, M. A. and Liang, P. (2023). "Foundation models and fair use," *arXiv preprint*, arXiv:2303.15715.
- Kizhner, I., Terras, M., Rumyantsev, M., Khokhlova, V., Demeshkova, E., Rudov, I., and Afanasieva, J. (2021). "Digital cultural colonialism: Measuring bias in aggregated digitized content held in Google Arts and Culture," *Digital Scholarship in the Humanities*, 36(3), pp. 607-640.
- Levendowski, A. (2018). "How copyright law can fix artificial intelligence's implicit bias problem," *Washington Law Review*, 93(2), pp.579-630.
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. (2022). "High-resolution image synthesis with latent diffusion models," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.10684-10695.
- Sandvik, K. (2011). "Book review," *MedieKultur: Journal of Media and Communication Research*, 50, pp.185-192.
- Shi, J., Nason, M., Tullney, M. and Alperin, J. P. (2025). "Identifying metadata quality issues across cultures," *College & Research Libraries*, 86(1), pp.1-30.
- Simon, N. (2010), *The participatory museum*, Santa Cruz, CA, Museum 2.0.
- Tanferri, M.(2022). "Digitizing heritage: More than clicking on a button, producing good copies," *New Sociological Perspectives*, 2(2), pp.94-108.
- Tao, Y., Viberg, O., Baker, R. S. and Kizilcec, R. F. (2024). "Cultural bias and cultural alignment of large language models," *PNAS Nexus*, 3(9), 346.
- Wu, R., Gong, C., Chen, L., Su, J. and Ray, L. C. (2024). "The present in the future is the past: Applying generative AI to visualize and imagine cultural heritage sites in both augmented and physical reality," *Proceedings of the 29th International Symposium on Electronic Art (ISEA)*, pp.824-831.
- Yusri, M. H. A., Johan, M. A, Khusaini, N. S, and Ramli, M. H. M. (2022). "Preservation of cultural heritage: A comparison study of 3D modelling between laser scanning, depth image, and photogrammetry methods," *Journal of Mechanical Engineering*, 19(2), pp.125-145.

- 저자 심정택은 서울시립대학교 경영대학 경영정보시스템 전공 박사과정에 재학 중이며, 국가유산진흥원 AI데이터팀 팀장으로 재직 중이다. 국립안동대학교 정보통신공학 학사, 숭실대학교 지식경영학 석사를 취득하였고, 숭실대학교 ITCEO과정과 KAIST AeG과정을 수료하였다. 주요 연구 분야는 국가유산 디지털 전환, AI 활용 전략, 소버린 AI 모델 개발 등이다.
- 저자 정재훈은 현재 서울시립대학교 경영대학 경영정보시스템 전공 조교수로 재직 중이다. 한국과학기술원(KAIST) 산업공학 학사, 서울대학교 대학원 협동과정 기술경영경제정책 전공 경영학 석사, 미국 메릴랜드대학교(University of Maryland) 정보시스템 (Information Systems) 전공 경영학박사를 취득하였다. 삼성전자, LG경제연구원, CJ주식회사 미래경영연구원, UNESCO 아태지역본부 등에 근무한 경험을 바탕으로 IT, 미디어/엔터테인먼트, 사회적 이슈를 아우르는 연구를 진행하고 있다.