

빅데이터와 인공지능 기반의 아동 그림책 이용자 요구 식별 및 디자인 연구

User Needs Identification and Design Strategies for Children's Picture Books Using Big Data and Artificial Intelligence

주 저 자 : 종 림 (Zhong, Lin)

한양대학교 디자인학부 시각디자인전공 박사과정

교 신 저 자 : 송민정 (Song, Min Jung)

한양대학교 디자인대학 커뮤니케이션디자인학과 교수
mjung111@hanmail.net

<https://doi.org/10.46248/kidrs.2025.3.497>

접수일 2025. 08. 17. / 심사완료일 2025. 08. 30. / 게재확정일 2025. 09. 08. / 게재일 2025. 09. 30.

Abstract

Despite the continuous decline in the domestic birth rate, the children's product market has been expanding. A wide range of picture books has been published for children's aesthetic enlightenment and cognitive development; however, their educational impact remains limited due to inconsistent content quality, prevalent homogenization, and the absence of systematic evaluation frameworks. In response, this study establishes a big data and AI-driven framework encompassing user needs identification, function mapping, and content validation, and systematically analyzes the content structure, modes of expression, and interaction of children's picture books. A total of 844,640 data entries were collected from 12 data sources, and semantic analysis combined with the TF-IDF method was employed to extract key user requirements and determine their priorities. These elements were then integrated with the reading process and user personas to identify critical problem points, which informed the construction of a functional tree for designing visual styles and plot interactions. Subsequently, prototype picture books were developed and tested in parent-child settings to evaluate readability, engagement, and educational effectiveness. Ultimately, this research proposes a data-driven design methodology that optimizes the content structure and representational strategies of children's picture books, providing a scholarly foundation for the systematic development of early childhood educational materials, while also suggesting new directions for personalized recommendation and structured evaluation systems.

Keyword

Big Data Analytics (빅데이터 분석), Artificial Intelligence (인공지능), Children's Picture Book Design (아동 그림책 디자인)

요약

최근 국내 출산율은 하락하고 있으나 아동 제품 소비시장은 확대되고 있다. 아동의 미적 계몽과 인지 개발을 위한 그림책은 다양하게 출간되고 있으나 콘텐츠 품질 편차와 동질화, 체계적 평가 시스템 부재로 교육 효과가 제한적이다. 이에 본 연구는 빅데이터와 인공지능을 기반으로 사용자 요구 발굴, 기능 매핑, 콘텐츠 검증까지 아우르는 프레임워크를 구축하여 그림책의 내용 구조, 표현 방식, 상호작용을 체계적으로 분석하였다. 총 12개 데이터 소스에서 844,640건의 데이터를 수집하고 의미 분석과 TF-IDF 기법을 통해 핵심 요구 요소와 우선순위를 도출하였다. 이를 독서 프로세스와 사용자 페르소나와 결합하여 핵심 문제 지점을 규명하고, 기능 트리를 기반으로 시각적 스타일과 플롯 상호작용을 설계하였다. 이후 샘플 그림책의 프로토타입 제작과 친자 테스트를 통해 가독성, 흥미도, 교육 효과를 검증하였다. 연구는 궁극적으로 아동 그림책의 내용 구조와 표현 양식을 데이터 기반으로 최적화하는 설계 방법론을 제시하며, 이는 조기교육 그림책의 체계적 개발에 학문적 근거를 제공하는 동시에, 개인화 추천과 구조화 평가 체계 구축을 위한 새로운 연구 방향을 제안한다.

목차

1. 서론

- 1-1. 연구 배경 및 목적
- 1-2. 연구 방법 및 범위

2. 이론적 배경

- 2-1. 빅데이터 사용자 요구 분석
- 2-2. 데이터 기반 인공지능
- 2-3. 아동 그림책과 행동 인지 심리

3. 사용자 요구 발굴 및 워드 클라우드 라이브리 구축

- 3-1. 다차원 데이터 수집 및 처리
- 3-2. 요구 의미 분석 및 시각화

4. 데이터 기반 아동 그림책 디자인 실습 및 효과 검증

- 4-1. 기존 그림책의 최적화 및 혁신 디자인 연구

- 4-2. 디자인 설명
- 4-3. 실험 프로세스 및 효과 검증

5. 결론

참고문헌

1. 서론

1-1. 연구 배경 및 목적

최근 아동 그림책은 조기교육과 아동문학의 중요한 매체로서 전 세계적으로 현저한 성장세를 보인다. Cognitive Market Research의 <Children Picture Book Market Report 2025> 조사 보고서에 따르면, 2024년 전 세계 아동 그림책 시장 규모는 약 47억 1,500만 달러로, 2031년까지 연평균 복합 성장률(compound annual growth rate, CAGR) 4.3%로 지속 성장하여 약 66억 달러에 달할 것으로 예상된다고 밝혔다. 북미 지역이 전 세계 시장 점유율의 약 40%를 차지하고 있으며, 아시아 태평양 지역은 23%의 시장 점유율과 6.3%의 높은 성장률로 가장 잠재력이 높은 성장 지역으로 평가되고 있다. 아동 조기 독서를 장기간 추진해 온 전통과 부모 돌 교육에 관한 관심이 높아진 가운데, 아태 지역(중국, 일본, 한국 등 포함)은 6.3%의 CAGR로 가장 빠른 성장세를 보이며, 2024년 시장 규모가 이미 10억 8,500만 달러에 달했다. 이러한 성장은 주로 전 세계 중산층 가정의 실소득 증가, 조기교육의 보편화, 다양성과 포용성을 담은 콘텐츠에 대한 수요 증가에 기인한다. 부모는 아동의 인지, 정서, 사회적 능력 발달을 촉진하기 위해 교육성과 오락성을 겸비한 그림책을 구매하는 경향이 점차 강해지고 있다.

한국의 아동 그림책 시장은 2024년 약 1억 800만 달러 규모로, 5.4%의 CAGR로 지속 성장할 것으로 예상된다. 한국 시장은 고도의 도시화와 디지털화 추세에 힘입어 부모 돌 조기교육 자원에 대한 수요가 왕성하며, 특히 영어 학습과 STEM(science, technology, engineering, mathematics) 관련 콘텐츠 분야에서 도시 거주 부모의 48%가 이러한 그림책을 선호하는 것

으로 나타났다. 또한 한국의 디지털 출판 플랫폼과 모바일 독서 애플리케이션의 보급이 그림책의 온라인 판매를 더욱 촉진하여, 온라인 소매가 전체 매출의 55%를 차지하고 있다. 그러나 한국의 아동 그림책 시장은 만화(manga) 등 대중 문화에 대한 선호로 인해 전통적인 그림책 시장 점유율이 상대적으로 감소하는 등의 과제에 직면해 있다.¹⁾

시장 성장 전망이 낙관적임에도 불구하고, 전 세계 및 한국의 아동 그림책 시장에는 여전히 몇 가지 문제점이 존재한다. 첫째, 그림책 품질이 천차만별이며 동질화 현상이 심각하여, 많은 제품이 혁신성과 과학적 근거가 부족해 아동의 감지 미학과 지능 개발 측면에서의 다양화된 요구를 충족하지 못하고 있다. 둘째, 생산비용 상승(시장 조사 및 디자인 비용 등)으로 인해 출판사들이 가격 경쟁력을 유지하기 어려워져 그림책의 접근성에 영향을 미치고 있다. 또한 디지털 엔터테인먼트(비디오 게임, 스트리밍 서비스 등)의 경쟁이 전통적인 그림책 시장에 압박을 가하여 아동의 독서 시간이 분산되고 있다. 한국의 경우 문화적 차이와 엄격한 출판 규제가 시장 진입과 확장에 어려움을 초래하고 있다. 동시에 다양성과 포용성을 담은 콘텐츠에 대한 수요가 증가하고 있음에도 불구하고, 시장에서 다문화와 다양한 배경을 진정으로 반영하는 양질의 콘텐츠는 여전히 부족한 상황이다.

따라서 본 연구는 빅데이터와 인공지능 도구를 기반으로 사용자 평가 행동 데이터에 대한 심층 분석을 통해 아동 그림책 디자인의 체계적 요구 모델을 구축하

1) Cognitive Market Research, Children Picture Book Market Report 2025, (2025.08.11), <https://www.cognitivemarketresearch.com/children-picture-book-market-report>.

고, 이러한 데이터를 실제 디자인 가치를 지닌 도문 콘텐츠, 상호작용 방식, 형식 구조로 전환하는 방법을 탐구하여 과학적이고 검증할 수 있는 아동 그림책 디자인 경로를 구축하고자 한다. 이를 통해 시장 구조 전환과 사용자 요구 개인화라는 이중 과제에 대응하고자 한다. 본 연구는 현재 디자인 이론과 방법론상의 공백을 메우는 현실적 의미를 지닐 뿐만 아니라, 향후 아동 도서 출판과 학제 간 융합을 위한 새로운 방향을 제시하는 데 목적이 있다.

1-2. 연구 방법 및 범위

학제 간 분석법. 아동 그림책 제품의 디자인 연구는 단순히 시각 디자인의 문제가 아니라, 빅데이터 기반 사용자 요구 분석을 통해 컴퓨터 과학, 데이터 사이언스, 아동 인지 심리학 등의 학문 분야를 포함한다. 따라서 사용자 요구 분석에 기반한 아동 그림책 디자인 연구는 연구 방법과 디자인 방안의 실현 가능성과 유효성을 보장하기 위해 학제 간 지식과 방법론이 필요하다.

정량·정성 결합 분석법. 본 연구에서 채택한 사용자 의미 추출, 빅데이터 마이닝, 아동 행동 데이터 분석 등은 정량 연구에 해당한다. 정성 연구는 주로 사용자 인터뷰, 사용자 페르소나, 사용자 행동 기록 등을 활용한다. 주관적·객관적 분석 방법을 결합한 아동 그림책 디자인 연구 지도를 통해 아동과 부모의 요구와 문제점을 더욱 포괄적으로 이해함으로써 아동 그림책의 디자인과 개발을 지도할 수 있다.

실험적 디자인 방법. 대표성과 실현 가능성을 갖춘 실물 모델 샘플을 제작하여 연구의 신뢰도와 정확성을 제고한다. 먼저 GAN 생성 적 적대 신경망 도구를 활용하여 아동 그림책 디자인 방안의 이미지와 효과도 렌더링을 완성한다. 둘째, 이론적 방안을 지침으로 삼아 아동 그림책 실물 샘플 모델을 제작한다. 마지막으로 샘플 모델을 통해 방안 검증한다.

2. 이론적 배경

2-1. 빅데이터 사용자 요구 분석

빅데이터 사용자 요구 분석은 빅데이터 기술을 활용하여 특정 주제나 제품에 대한 네트워크 리뷰를 체계적으로 분석하는 것으로, 주로 사용자 요구 획득, 모델링, 전환이라는 세 단계를 포함한다. 의미 추출을 통해

단시간 내에 사용자가 자발적으로 작성한 대량의 제품 평 텍스트를 획득할 수 있으며, 사용자 요구 모델링을 통해 사용자 요구 간의 특성 관계를 정확히 정리할 수 있다. 최종적으로 사용자 요구 전환 기술을 활용하여 사용자 요구를 제품 기능과 해결 방안으로 대응함으로써 제품의 사용자 요구에 대한 응답 효율성과 정확도를 제고한다. 의미 추출 과정에서 핵심 기술은 자연어 처리(natural language processing, NLP), 텍스트 마이닝, 데이터 마이닝 등을 포함한다. 해외 연구 측면에서 Quan 등(2023)은 빅데이터와 AI의 결합이 다중 모달 데이터(텍스트, 이미지, 오디오 등) 처리를 통해 제품 디자인 효율성을 현저히 향상하게 시키고 전통적 방법의 한계를 극복한다고 지적하며, 향후 데이터 융합 기술에 집중해야 한다고 제시하였다. 2) Mortati 등은 8개 사례 분석을 통해 빅데이터(정량)와 두꺼운 데이터(정성)가 디자인 사고에서 협력하여 복잡한 문제에 효과적으로 대응하고 혁신을 추진한다고 강조하였다. 3) Schimpf와 Goldstein(2022)은 LaDDR 프레임워크를 제시하여 빅데이터를 활용해 디자인 활동 데이터를 포착하고 디자인 연구의 확장성을 향상하게 시켰다. 4) Fahmideh와 Beydoun(2019)은 빅데이터 분석 아키텍처를 제안하여 제조 시스템의 제품 생명주기 관리를 최적화하고 데이터 기반 의사결정 능력을 강화하였다. 5) 국내 연구 측면에서 서양은 동적 그래픽 로고의 시각적 요소가 빅데이터 분석을 통해 사용자의 감정 반응을 식별할 수 있어 감성적 브랜드 디자인에 방향을 제공한다고 제시하였으며, 6) 윤수인 등(2022)은 AI

- 2) Huafeng Quan, Shaobo Li, Changchang Zeng, Hongjing Wei, Jianjun Hu, 'Big Data and AI-Driven Product Design: A Survey', Applied Sciences, 2023. 01, Vol.13, No.16, pp. 94.
- 3) Marzia Mortati, Stefano Magistretti, Cabirio Cautela, Claudio Dell'Era, 'Data in design: How big data and thick data inform design thinking projects', Technovation, 2023. 04, Vol.122, pp. 102.
- 4) Corey Schimpf, Molly H. Goldstein, 'Large data for design research: An educational technology framework for studying design activity using a big data approach', Frontiers in Manufacturing Technology, 2022. 10, Vol. 2.
- 5) Mahdi Fahmideh, Ghassan Beydoun, 'Big data analytics architecture design—An application in manufacturing systems', Computers & Industrial Engineering, 2019. 02, Vol. 128, pp. 948–963.
- 6) 서양, 신인식, '소셜 미디어 빅데이터 분석을 활용한 모션그래픽 로고의 시각적 감성 분석', 조형 미디어학, 2023. 02, Vol. 26, No.1, pp. 232–239.

패션 플랫폼 연구에서 색상 일치성과 이미지 강조성이 사용자 사용성을 현저히 향상하게 시킨다고 지적하였다.⁷⁾ 성광숙(2021)은 소셜 미디어 데이터를 바탕으로 Z세대의 패션 분야 핵심 관심 키워드를 밝혔고, ⁸⁾남미우(2020)는 Instagram 메이크업 게시물을 분석하여 성별 미적 표현에 현저한 차이가 있음을 발견하였다.⁹⁾ 장진희(2020)는 보석 시장이 판매자 중심에서 데이터 기반 사용자 소통으로의 전환이 시급하다고 강조하였다.¹⁰⁾

따라서 선행연구들은 빅데이터 활용이 디자인 연구 및 제품 개발의 효율성을 높이는 데 기여함을 보여주었지만, 여전히 몇 가지 한계가 존재한다. 기존 연구가 특정 산업(패션, 제조, 마케팅 등)에 치중되어 있어 아동 그림책과 같은 교육, 문화 콘텐츠 영역에서는 적용 사례가 부족하다. 또한 대부분의 연구가 정량적 데이터 분석에 집중되어 정성적 맥락을 충분히 반영하지 못한다는 점에서 보완이 필요하다. 본 연구는 이러한 간극을 메우기 위해, 아동 발달과 학습 환경을 고려한 사용자 요구 분석을 시도하고 데이터 기반 설계 방법론을 그림책 디자인에 접목함으로써 학문적 확장 가능성을 모색하고자 한다.

2-2. 데이터 기반 인공지능

데이터 기반 방법은 대량의 기존 데이터를 수집하고 인공지능 알고리즘을 활용한 학습을 통해 기존 데이터에서 디자인 규칙을 학습하며, 수동으로 규칙을 정의하지 않고 데이터에서 학습하므로 높은 범용성과 적응성을 갖는다. Cui 등 (2019)은 텍스트 생성 인포그래픽 모델을 설계하여 비주얼 데이터를 미리 설정된 스타일의 인포그래픽으로 자동 전환할 수 있도록 하였다.¹¹⁾ Li 등 (2020)은 차별화할 수 있는 렌더링, 조

건부 GAN 등의 모델을 제안하여 레이아웃 정렬과 국부 구조 학습을 실현하였다.¹²⁾ Hsu 등 (2023)은 인간 디자인 과정을 모방하여 레이아웃을 재구성하였고,¹³⁾ Zhou 등 (2022)은 시각 그래픽 레이아웃 GAN을 제안하여 이미지 내용과 결합한 공간 구조를 생성하였다.¹⁴⁾ Li 등[20]은 관계 인식 확산 모델을 설계하여 시각과 텍스트 정보를 통합하였다.

황윤정 (2023) 연구는 생성형 AI(예: Midjourney)가 브랜드 아이덴티티 디자인을 효율적으로 창조할 수 있지만, 그 결과물은 디자이너의 개념에 대한 정확한 이해와 키워드 입력에 의존하므로 인력을 대체하기보다는 보조 도구로 활용되어야 한다고 지적하였다.¹⁵⁾ 최효승 (2017)은 AI 예술이 데이터 재조합의 창조성, 오락성, 편의성, 가변성을 나타내지만, 인간의 직관적 창조력을 대체할 수는 없으며, 그 의미는 과학기술과 예술의 학제 간 융합을 추진하는 데 있다고 보았다.¹⁶⁾ 정은희는 AI 도구가 디자인 프로세스의 창의 생성 단계를 지원할 수 있지만 기획과 의사결정 보

11) Weiwei Cui, Xiaoyu Zhang, Yun Wang, He Huang, Bei Chen, Lei Fang, Haidong Zhang, Jian-Guan Lou, Dongmei Zhang, 'Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements', arXiv, 2019, <http://arxiv.org/abs/1907.09091>.

12) Jianan Li, Jimei Yang, Jianming Zhang, Chang Liu, Christina Wang, Tingfa Xu, 'Attribute-conditioned Layout GAN for Automatic Graphic Design', arXiv, 2020, <http://arxiv.org/abs/2009.05284>.

13) HsiaoYuan Hsu, Xiangteng He, Yuxin Peng, Hao Kong, Qing Zhang, 'PosterLayout: A New Benchmark and Approach for Content-Aware Visual-Textual Presentation Layout', '2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)', IEEE, 2023, pp.6018-6026.

14) Min Zhou, Chenchen Xu, Ye Ma, Tiezheng Ge, Yuning Jiang, Weiwei Xu, 'Composition-aware Graphic Layout GAN for Visual-Textual Presentation Designs', 2022, pp.4995-5001.

15) 황윤정, 'BI 개발을 위한 생성형 인공지능 디자인 플랫폼 활용에 관한 연구 - '미드저니'의 활용 가능성을 중심으로-', 한국디자인문화학회지, 2023. 09, Vol. 29, No. 3, pp.527-541.

16) 최효승, 손영미, '인공지능과 예술 창작 활동의 융복합 사례 분석 및 특성 연구', 한국 과학 예술 융합 학회, 2017. 03, Vol. 28, pp.289-299.

7) 윤수인, 한정민, 박혜린, 김지인, '트렌드 빅데이터 분석 기반의 AI 패션 플랫폼 UX/UI 디자인 개발', 아시아태평양 융합 연구 교류 논문지, 2024. 07, Vol. 10, No. 7, pp. 97-110.

8) 성광숙, 'Z세대 패션에 대한 소셜 미디어의 빅데이터 분석', 한국의상디자인학회지, 2020. 09, Vol. 22, No. 3, pp. 49-61.

9) 남미우, '빅데이터 분석 방법을 적용한 인스타그램 미용 관련 해시태그 게시물 분석에 관한 융합 연구', 한국 과학 예술 융합 학회, 2019. 09, Vol. 37, No. 4, pp. 93-106.

10) 장진희, '빅데이터를 활용한 주얼리 마케팅 제안', 한국 과학 예술 융합 학회, 2015. 03, Vol. 19, pp. 593-600.

조 측면에서는 여전히 부족하다고 발견하였다.¹⁷⁾

기존 연구는 주로 생성형 AI의 기술적 효율성에 집중하였으나, 아동 발달 단계나 정서적 맥락을 충분히 반영하지 못하는 한계가 있다. 본 연구는 데이터 기반 자동화를 아동 교육·심리적 특성과 결합함으로써, AI를 단순한 보조 도구가 아닌 아동 그림책 디자인의 새로운 패러다임으로 확장하는 데 의의를 둔다.

2-3. 아동 그림책과 행동 인지 심리학

최근 아동 그림책의 행동인지 심리학 분야 응용 연구가 주목받고 있다. Ganea 등(2008)은 15~18개월 영아가 고도 구성 그림책(실제 사진 등)을 통해 라벨을 현실 물체로 전이할 수 있음을 발견하여, 그림책이 초기 기호 발달과 행동인지에서 갖는 역할을 밝혔다.¹⁸⁾ Strouse 등(2017)은 그림책 형식(전자 vs. 인쇄 등)과 특성(실재성, 애니메이션 등)이 친자 상호작용, 언어 입력, 유아의 지식 전이에 현저한 영향을 미친다고 더욱 지적하였으며, 애니메이션 그림책이 4~6세 아동의 시각적 주의력과 이야기 이해를 향상하게 시킬 수 있다고 밝혔다.¹⁹⁾ Montag 등(2015)은 어휘 통계 분석을 통해 그림책이 다양화된 어휘를 제공하여 언어와 인지 발달을 촉진한다고 강조하였다.²⁰⁾

기존 연구들은 빅데이터와 인공지능을 활용한 디자인 방법론의 효율성을 보여주었고, 아동 그림책이 초기 인지 발달과 언어 학습에 중요한 역할을 한다는 사실도 입증해 왔다. 그러나 두 영역은 여전히 분절적으로 다루어지고 있어, 아동 독서 경험을 정량적 데이터 분석과 긴밀히 연계하는 방안은 충분히 논의되지 못하였다. 이러한 점은 그림책이 지녀야 할 교육적 효과와 정

서적 경험의 균형을 확보하는 데 한계를 남긴다. 따라서 데이터 기반 분석을 아동 발달 연구와 결합하여 그림책 디자인의 최적화 가능성을 탐색하는 작업은 향후 보다 심화한 논의가 요구되는 부분이라 할 수 있다.

3. 사용자 요구 발굴 및 워드 클라우드 라이브러리 구축

본 연구의 데이터 선정 범위는 아동 그림책 관련 사용자 데이터에 집중하며, 공공 데이터베이스(Amazon Review Data, 그림책 박물관 등), 문화 분야 빅데이터 플랫폼(good reads book graph data sets, recbole dataset list), 고인이기 전자상거래 플랫폼(Coupang, g market 등), 다수의 아동 그림책 브랜드 및 기업 웹사이트(YES24, 가온 빛, 한국 프로젝트 출판사 등) 등 총 12개 주요 데이터 소스를 포함한다. 수집 내용은 그림책 명칭, 콘텐츠 소개, 사용 나이, 사용자 리뷰, 사용 피드백 등의 한국어나 영어 텍스트를 포함하여 광범위하고 대표성을 갖도록 보장하였다. 데이터 유형은 사용자 생성 콘텐츠(UGC)를 중심으로 하며, 도서 메타데이터 설명 텍스트로 보완하고, Python 크롤링과 구조화 전처리를 통해 최종적으로 표준화된 코퍼스(총 844,640건의 기록)를 형성하여 후속 분석 모형화에 활용하였다.

해당 데이터 범위를 선택한 주요 이유는 높은 실용성과 실제 피드백 가치를 갖기 때문이다. 한편으로는 공공 데이터베이스와 전자상거래 플랫폼의 리뷰가 실제 사용 상황에서의 사용자 선호도, 문제점, 요구사항을 반영할 수 있어 분석의 유효성을 강화하며, 다른 한편으로는 다양한 브랜드와 플랫폼의 텍스트 차별성이 다양한 디자인 언어와 주제 키워드 발굴에 도움이 된다. 한국 현지 데이터와 해외 데이터를 결합한 것은 아동 교육, 미적 습관, 문화 교류 측면에서 대표성을 확보하기 위함이다. 이러한 데이터 선택은 연구 결과의 일반화 가능성을 높이며, 학문적으로도 유의미한 시사점을 제공한다. 전체적으로 이러한 데이터는 워드 클라우드 모델링과 키워드 발굴에 적합할 뿐만 아니라, 후속 아동 그림책 디자인과 기능 매칭을 위한 고품질 데이터 지원을 제공한다.

3-1. 다차원 데이터 수집 및 처리

본 장에서는 아동 그림책 디자인을 위한 워드 클라

17) 정은희, 최정민, 'AI 디자인툴의 디자인 업무 지원 방향성 제안', Archives of Design Research, 2022. 11, Vol. 35, No. 4, pp.269-282.

18) Patricia A. Ganea, Megan Bloom Pickard, Judy S. DeLoache, 'Transfer between Picture Books and the Real World by Very Young Children', Journal of Cognition and Development, 2008. 02.

19) Gabrielle A. Strouse, Patricia A. Ganea, 'Parent-Toddler Behavior and Language Differ When Reading Electronic and Print Picture Books', Frontiers in Psychology, 2017. 05, Vol. 8.

20) Jessica L. Montag, Michael N. Jones, Linda B. Smith, 'The Words Children Hear', Sage Journals, 2015. 08, Vol. 26, No. 9, pp1489-1496.

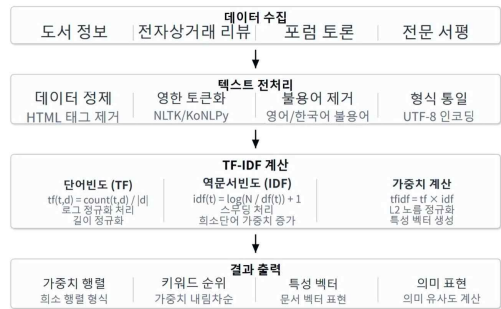
우드 라이브러리를 구축하고 적용하여, 네트워크 빅데이터 수집, 텍스트 전처리, 키워드 추출, 시각화를 통해 최종적으로 아동 그림책 디자인의 핵심 특징과 최적화 동향을 분석하고 규명한다.

먼저 크롤링 프로그램을 작성하여(python의 requests와 beautiful soup을 채택하여 정적 페이지를 파싱하고, 필요시 Parse Hub 시각화 웹 데이터 추출과 결합) 그림책 명칭, 콘텐츠 소개, 목표 연령대, 사용자 평가, 사용 피드백 등의 텍스트 콘텐츠를 자동으로 추출하고, 결과를 UTF-8 인코딩의 CSV 형식으로 구조화하여 저장함으로써 후속 분석 사용에 편의를 제공한다.

[표 1] 데이터 출처와 유형

데이터 출처	웹사이트 주소	데이터 형태
Amazon Review Data (2023)	https://huggingface.co/datasets/McAuley-Lab/Amazon-Reviews-2023	리뷰 데이터
Recbole Data-set List	https://recbole.io/dataset_list.html	리뷰 데이터
The Horn Book Guide	https://www.hornbookguide.com/site/?page=about-the-horn-book-guide&utm_source=com	상품 소개 및 사용자 리뷰
Children's Picture Book Database at Miami University	https://dpl.lib.miamioh.edu/picturebook/?utm_source=com	도서 메타데이터
Goodreads Book Graph Datasets	https://cseweb.ucsd.edu/~jmcauley/datasets/goodreads.html	사용자 리뷰
그림책 박물관	https://www.picturebook-museum.com/main	도서 메타데이터
Yes24	https://www.yes24.com/Mall/School/ReadingGuidance	사용자 리뷰
어린이도서 연구회	https://www.childbook.org/#5rdPage	상품 소개 및 사용자 리뷰
Coupang	https://www.coupang.com/	사용자 리뷰
한국 프로젝트 출판사	https://project21.com/main.php	도서 메타데이터 및 사용자 리뷰
G market	http://gcategory.gmarket.co.kr	사용자 리뷰
가온빛	http://gaonbit.kr/feature/98263/	도서 메타데이터

원시 데이터 수집 완료 후, 그림책 콘텐츠와 특징 설명을 포함하는 ‘설명’ 필드에 대해 텍스트 전처리를 시행한다. 한국어 텍스트 처리는 형태 변화와 입도 문제를 고려해야 한다. 기본 단계는 잡음 제거, 형태소



[그림 1] TF-IDF 방법 기술 로드맵

분석 및 품사 태깅, 불용어 필터링, 어간 추출을 포함한다. 전처리 과정은 먼저 HTML 잔여 태그, 스크립트, 광고 등의 불필요한 요소를 제거한 후, 한국어 형태소 분석 도구(KoNLPy 프레임워크의 Mecab, Okt 등)를 통해 각 기록에 대해 형태소 분석과 품사 태깅을 시행하여 어간을 추출하고 동사, 형용사를 기본형으로 복원하며, 완전한 명사 형태를 보존한다. 예를 들어, 한국어 단어 ‘결정하겠따’라는 ‘결정(결정) + 하(하다) + 겠(추측의지) + 다(어미)’로 분해되며, 형태 분석을 통해 어간과 품사를 추출할 수 있어 후속 키워드 필터링에 도움이 된다. 이후 조사(‘이/가/은/는/을’), 접속사 및 기타 기능어를 필터링하고, 사용자 정의 불용어 목록을 활용하여 고빈도이지만 실질적 정보가 없는 단어와 상투어를 제거하며, 대소문자를 통일하고(영문자가 혼재할 경우), 숫자와 단위 형식을 표준화하고, 저빈도 또는 고빈도로 정보량이 부족한 단어(연속어 등) 늘 제거한다. 코퍼스 규모에 따라 정규표현식이나 사전을 사용하여 동의어를 치환한다. 처리 후 순수한 한국어 단어 목록을 얻어 후속 분석과 모델링에 활용할 수 있다.

3-2. 요구 의미 분석 및 시각화

키워드 추출 단계에서는 TF-IDF, 주제 모델 방법을 채택한다.

TF-IDF는 모든 문서에서 출현 빈도가 높으면서 다른 문서에서는 상대적으로 적게 나타나는 단어를 식별하여 가장 대표적인 디자인 요소를 추출하는 데 사용된다.

[표 2] TF-IDF 방법의 장점

방법	설명	적용 상황
TF-IDF	단어 빈도(TF)와 역문서 빈도(IDF)를 기반으로 단어 가중치를 계산하며, 출현 빈도가 높으면서 다른 문서에서는 상대적으로 적게 나타나는 단어에 높은 가중치를 부여한다. 특정 평가 집합이나 설명 집합에서 가	각 그림책 범주나 평가 집합별로 개별 계산하여 공통 주제어 추출을 가능하게 하는 방법

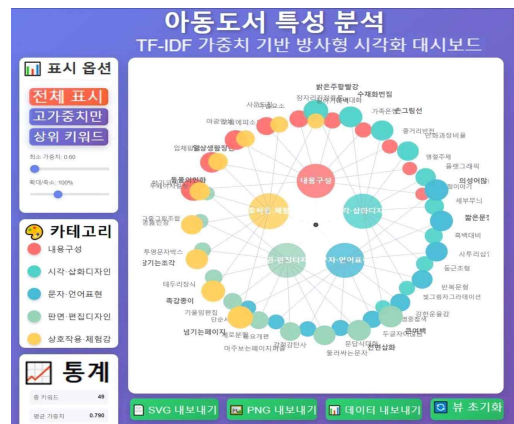
	장 대표적인 키워드를 추출하는 데 적합하다.	
--	--------------------------	--

[표 3] 가중치 상위 50개 키워드

범주	키워드	가중치	범주	키워드	가중치
내용구성	동물 의인화	0.156	판면및편집디자인	강한 운율감	0.119
	일상생활 장면	0.142		이음절어 다수	0.106
	모험 에피소드	0.138		문답식 대화	0.094
	마법 요소	0.124		감정 감탄사	0.083
	취침 안정 플롯	0.117		동요 개편	0.072
	유머러스한 대화	0.109		단순 서술	0.061
시각및일러스트레이션디자인	가족 따뜻함	0.098		큰 여백	0.163
	플롯 반전	0.085		전면 삽화	0.149
	명절 이야기	0.076		둘러싼 텍스트	0.137
	우정 이야기	0.063		대면 퍼즐	0.126
	밝고 따뜻한 색상	0.168		새로 단 나누기	0.115
	수채화 번짐	0.151		가을인 편집	0.103
	손 그림 선	0.143		테두리 장식	0.091
	카툰 과장 비율	0.132	상호작용및체험감	투명 텍스트 갑	0.079
	플랫 그래픽	0.125		다중 이미지 조합	0.068
텍스트및언어표현	세부 패턴	0.114		양면 확장	0.057
	흑백 대비	0.102		넘기는 페이지	0.174
	둥근 조형	0.089		축감하지	0.161
	광영 그래데이션	0.077		당기는 조각	0.148
	투명 중첩색	0.065		뚫린 창	0.135
	의성어 다수	0.159		향기 코팅	0.122
	짧은 문장	0.147		입체 팝업	0.108
	방언 삽입	0.136		야광 인쇄	0.095
	반복 문형	0.128		사운드 칩	0.082

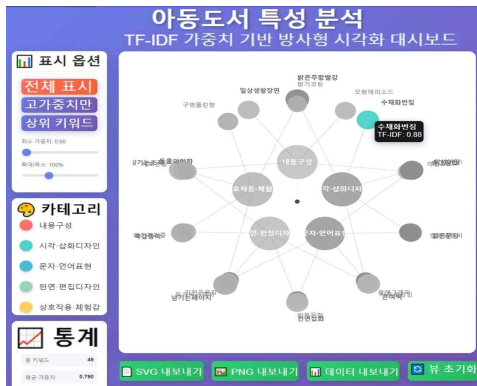
844,640건 기록의 정제된 텍스트를 코퍼스로 활용하여, 각 기록의 형태소 분석 결과에 대해 가중치를 계산하고, 가중치 순으로 정렬한 후 상위 50개 키워드를 선정하여 시각화 생성을 진행한다(오른쪽<표 3>과 같음).

아동 그림책 특성의 TF-IDF 가중치 분포를 보다 직관적으로 보여주기 위해, 본 연구는 D3.js 시각화 라이브러리를 활용하여 상호작용 형 방식형 워드 클라우드 시각화 시스템을 구축하였다. 시스템은 먼저 TF-IDF 처리 후의 키워드와 그 가중치 데이터를 읽어드리고, 방식형 레이아웃 디자인을 채택하여 캔버스 중심을 원점으로 하여 다섯 가지 주요 특성 범주(내용 구성, 시각 및 일러스트레이션 디자인, 텍스트 및 언어 표현, 판면 및 편집 디자인, 상호작용 및 체험감) 늘 원주상에 등각도로 분포시켜 위계화된 정보 아키텍처를 형성한다.



[그림 2] 1급 워드 클라우드 인터페이스

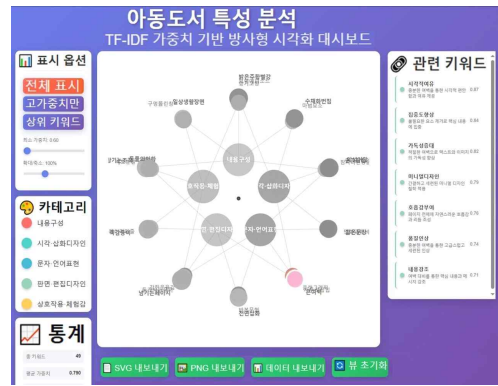
<그림 2>와 같이, 시각적 인코딩 측면에서 시스템은 다차원 매핑 전략을 채택하여 데이터 특성을 표현한다. 키워드의 중요도는 원의 크기를 통해 직관적으로 구현되며, 그 중의 원의 반지름은 가중치 제곱 매핑 공식을 사용하여 계산하고($r = \text{기초 반지름} + \text{가중치}^2 \times \text{확대 계수}$), 고 가중치와 저 가중치 어휘 간의 시각적 대비를 강화한다. 동시에 시스템은 각 특성 범주에 고유한 색상 식별을 할당하고, 가중치에 따라 서로 다른 면적과 색상을 부여하며, 고채도 배색 방안을 채택하여 서로 다른 범주 간의 명확한 구분을 보장한다. 글꼴 크기 역시 가중치 값과 정비례하며, 가중치가 0.085를 초과하는 고빈도 어휘는 굵은 글씨로 표시하여 그 중요성을 부각하게 시킨다.



[그림 3] 2급 워드 클라우드 인터페이스

<그림 3>과 같이, 시스템의 상호작용 기능 디자인은 사용자의 분석 요구를 충분히 고려하였다. 사용자가 인터페이스에서 임의의 키워드를 클릭하면 on('click') 이벤트가 트리거되어 투의 팁으로 해당 어휘의 상세 정보(구체적인 가중치 수치와 소속 범주 포함)를 보여준다. 좌측 제어 패널은 유연한 데이터 필터링 메커니즘을 제공하며, 사용자는 가중치 임계값 슬라이더(0.6~1.0 범위)를 통해 표시 내용을 동적으로 필터링하거나, '고 가중치 어휘와 '간소 보기' 같은 사전 설정된 표시 모드를 선택할 수 있다. 범주 범례는 클릭 전환 기능을 지원하여 사용자가 특정 특성 차원에 집중하여 심층 분석할 수 있도록 한다. 또한 범주 간 대비 분석을 지원하기 위해, 시스템은 사용자가 범주 필터링 기능을 통해 서로 다른 차원의 키워드 분포를 나란히 표시할 수 있도록 하며, 애니메이션 형태로 데이터의 변화 추세를 동적으로 보여줄 수 있다.

시스템은 또한 확대/축소 제어(50%~200%)와 다중 형식 내보내기 기능(SVG, PNG, JSON)을 통합하여 다양한 응용 시나리오의 요구를 충족한다(<그림 4>와 같음). 우측 통계 패널은 현재 뷰의 핵심 지표(어휘 총수, 평균 가중치, 최대 가중치 값 포함)를 실시간으로 업데이트하여 경량 분석을 위한 데이터 지원을 제공한다. 간소화된 위계 구조 개념도를 통해 연관 지도는 중심 워드 클라우드 노드에서 해당 모듈로 연결되는 전환 경로를 직관적으로 제시하여, 후속 기능 반복과 인터페이스 최적화를 위한 시각화 의사결정 근거를 제공한다.



[그림 4] 3급 워드 클라우드 인터페이스

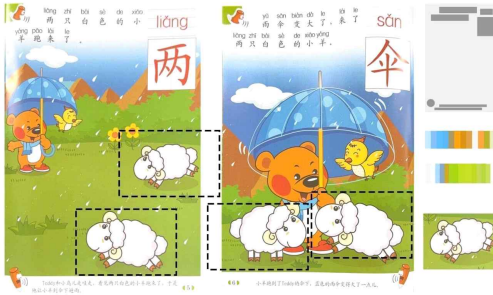
4. 데이터 기반 아동 그림책 디자인 실험 및 효과 검증

본 연구는 데이터 기반 설계 방법론을 핵심으로 하여, 그림책의 내용 구조 최적화, 시각 및 삽화 재구성, 언어 및 지면 구성 설계, 상호작용 경험 강화의 네 가지 측면에서 실천을 전개하였다. 기존 그림책의 재 디자인과 주제 지향형 오리지널 그림책 창작을 병행하고, 인공지능 기반 삽화 제작과 키워드 단어 구름 분석을 결합하여 검증할 수 있고 재활용할 수 있는 아동 그림책 설계 프로세스를 구축하였다. 나아가 현장 실험을 통해 해당 설계의 실효성과 적용 가능성을 체계적으로 검증하였다.

4-1. 기존 그림책의 최적화 및 혁신 디자인 연구

본 연구에서는 제안한 설계 방법의 실효성을 검증하기 위하여 두 유형의 그림책을 실험 대상으로 선정하였다.

길림 출판사에서 발간한 <Teddy의 우산>(그림 5)을 사례로 하였으며, 본 도서는 2~7세 아동을 대상으로, 곰 인형 Teddy와 크기 변화가 가능한 신기한 우산을 소재로 한 이야기로, 스토리-아동 한자 학습-멀티미디어 점도가 기능을 결합한 복합형 그림책이다.본 연구에서는 해당 그림책을 데이터 기반 설계 방법론의 실험 사례로 활용하였다.



[그림 5] Teddy의 우산 현존 문제 분석

그러나 질적 분석과 사용자 인터뷰 결과, 원본 <Teddy의 우산>은 시각 표현 및 사용자 경험 측면에서 다방면의 한계가 확인되었다. 우선, 삽화 스타일은 캐릭터 형태가 단순화된 카툰에 치우쳐 세부 묘사가 부족하고 표정이 단일하여 풍부한 감정 전달이 어려웠다. 배경 환경은 공간적 층위감과 분위기 연출이 부족하며, 주택·잔디 등의 요소가 평면적으로 표현되었고 빗방울 형태는 기호화가 두드러져 장면의 사실성과 역동성이 약화되었다. 색채 활용에서는 채도가 과도하게 높고 배경과 캐릭터 색상 대비가 강하나 전환이 부자연스러우며, 전체 색채와 선의 스타일이 통일되지 않아 시각적 분절감을 유발했다. 지면 구성 측면에서는 텍스트와 그림의 결합도 낮아 배치와 삽화가 분리되어 보였고, 원할한 독서 유도선이 부족했다. 서체는 성체를 사용하여 선이 가늘고 재미 요소가 부족해 2~7세 아동의 신속한 인지와 주의 집중에 불리했으며, 병기된 병음과 한자의 제시 방식이 단조롭고 흥미성이 결여되었으며, 점도가 인터랙션 영역의 표시가 불명확해 오작동 또는 인식 누락이 발생하기 쉬웠다.

‘뽀뽀 혼자 자기’를 주제로, 2~6세 아동의 독립 수면 능력 함양을 목표로 하였다. 앞서 구축한 단어 구름(워드 클라우드) 키워드 데이터베이스(예: ‘동물 의인화’, ‘밝고 따뜻한 색채’, ‘반복 문항’, ‘플랩북’)를 결합하여 스토리 주제, 플롯 구성, 색채 기조, 지면 구조 및 상호작용 형식을 제로베이스에서 설계하였으며, 모든 설계 의사결정이 데이터 분석 결과에 근거하도록 하였다.

두 사례 모두에서, TF-IDF 기반의 키워드 가중치 창작 참조 체계를 적용하여 고 가중치 어휘를 실행할 수 있는 시각 및 서사 지침으로 변환하였다. 예를 들어, 시각 디자인에서는 ‘부드러운 냉온 대비’를 주조 색으로 설정하고, ‘둥근 형태’를 캐릭터 선행 규범으로, ‘수

채 번짐’ 기법을 배경 질감 처리안으로 반영하였다. 삽화 제작 과정에서는 생성형 AI 그림 도구에 선별된 키워드 조합을 입력하여 다중 버전 스케치를 생성한 후, 인공 지검과 세부 보정을 거쳐 시각적 시안의 효율적 생성과 스타일 통일성을 확보하였다. 지면 구성 및 상호작용 설계 단계에서는 사용성 테스트와 저연령 독서 습관 연구를 바탕으로 텍스트 영역 비율을 조정하고, 고가 독성의 둥근 체 서체를 적용하며, 상호작용 핫스팟 영역을 확대하여 정확성과 흥미성을 향상하게 시켰다.

4-2. 디자인 설명

재 디자인된 <Teddy의 우산>(그림 6)에서는 빅데이터 단어 구름의 의미 분석 결과를 기반으로 스토리 상황 키워드, 아동 감정 선호 어휘, 환경 묘사 요소를 정밀 추출·가중치 화하여, 이들 고빈도·고 연관도의 시각 기호를 삽화 제작에 반영하였다. 삽화 스타일은 손 그림 질감의 색연필 풍으로 전환되어 부드러운 필치와 자연스러운 색채 전환을 통해 전반적으로 따뜻하고 친



[그림 6] 최적화된 버전



[그림 7] 디자인 전후 비교

근한 분위기를 형성하였다.

예를 들어 그림 7의 표기 1에서는 캐릭터의 외형·표정·신체 언어를 세밀하게 묘사하여 두 뺨 홍조·미소 등을 통해 생동감 있는 감정 전달이 가능하게 하였다. 표기 2에서는 단어 구름 분석에서 도출된 고빈도 자연 요소(꽃, 녹색 식물, 빗방울 등)를 배경에 추가하여 몰입감과 공간 충위를 강화하였다. 색채 활용은 채도를 낮추고, 캐릭터와 배경의 색조·선 굵기·스타일을 통일시켜 시각적 연속성과 안정감을 높였다. 지면 구성은 텍스트와 그림의 유기적 결합을 통해 자연스러운 시선 유도를 실현하였으며, 표기 3에서처럼 굵고 둥글며 아동 친화적인 서체를 사용하여 가독성과 흥미를 향상하게 시켰다. 표기 4에서는 한자와 병음을 창의적으로 화면에 배치(예: 'yǔ 雨'를 왼쪽 아래 배치) 하여 미감과 교육 기능을 동시에 구현하였다. 표기 5의 점도가 표시는 화면과의 시각적 융합이 높아 기능성을 유지하면서도 사용 편의성을 향상하게 시켜 독서 경험·정서 몰입·참여도를 크게 개선하였다.



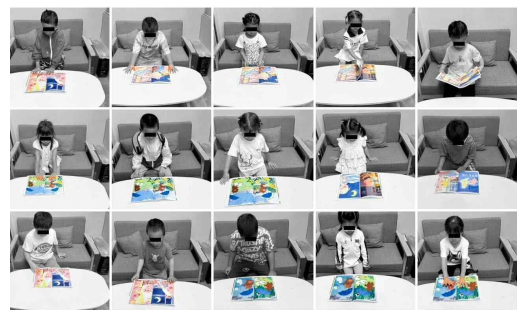
[그림 8] 데이터 기반의 혁신적 그림책 디자인

〈뽀뽀 혼자 자기〉(그림 8)의 창작 과정에서는 ‘독립 수면’ 주제로 아동이 자주 언급한 고빈도 어휘(‘용감함’, ‘안전함’, ‘동반’, ‘달’, ‘따뜻함’ 등)를 그림책의 플롯과 세부 장면에 적극 반영하였다. 표지와 주요 장면에서는 ‘용감한 작은 전사의 상징 이미지를 강화하고, 부드럽고 따뜻한 색조의 야간 장면을 통해 심리적 안정감을 조성하였다. 캐릭터의 표정과 동작 설계는 단어 구름의 감정 연관어의 정서 경향을 참고하여 긴장에서 이완, 의존에서 독립으로의 심리 변화를 시각적으로 구현하였다. 텍스트는 고빈도 어휘의 의미 네트워크를 바탕으로 대사를 최적화하여 아동의 언어 습관에 맞추고 친근감을 높였다. 예를 들어, ‘작은 전사’, ‘수호 보물’ 등의 호칭은 단어 빈도 분석 결과에서 도출되었으며, 감정 어

휘 조합을 통해 따뜻한 심리적 암시 체계를 형성하였다. 전체 설계는 데이터 기반과 감성 서사의 융합을 통해 아동이 몰입형 독서 과정에서 ‘혼자 자기’에 대한 긍정적 암시와 모방 동기를 자연스럽게 수용하도록 하는 것을 목표로 하였다.

4-3. 실험 프로세스 및 효과 검증

데이터 기반 그림책 디자인 방법의 유효성을 검증하기 위해, 본 연구는 목표 연령대 아동 집단을 대상으로 2주간의 현장 실험을 진행하였다. 실험 대상은 동일 유치원 출신 2~7세 아동 48명으로, 실험군과 대조군으로 무작위 배정하여 각 그룹당 24명으로 구성하였다. 실험 자료는 재 디자인된 〈Teddy의 우산〉과 창작 그림책〈뽀뽀 혼자 자기〉, 그리고 실험에서는 원본 그림책과 데이터 기반으로 재 설계된 그림책을 함께 제시하여 두 버전 간의 효과를 비교할 수 있도록 하였다. 실험 과정은 교실과 독서 코너의 이중 환경에서 진행되었으며, 교사와 학부모가 공동으로 협력하여 완성하였다.



[그림 9] 실험 현장 사진(일부)

먼저 연구팀은 참여자의 독서 연습관, 문해 수준, 주의력 지속시간에 대한 사전 평가하여 실험 기준선의 비교 가능성을 보장하였다. 이후 2주 주기 동안 각 아동은 매주 2회의 그림책 독서 체험을 받았으며, 독서 과정은 모두 동일한 시간대와 동일한 환경 조명 조건에서 진행되었고, 비디오 기록과 상호작용 로그 수집 시스템을 통해 독서 행동 데이터와 정서 반응을 동시에 기록하였다. 독서 종료 후 아동은 단서 제공 없이 이야기 줄거리를 복술 해야 했으며, 간단한 상호작용 문답을 진행하였고, 학부모와 교사는 각각 독서 만족도와 상호작용 적극성 평가표를 작성하였다.

[표 4] 아동 그림책 읽기 실험 결과

지표	실험군	대조군	차이(%)	p 값
평균 독서 지속시간(분)	12.8 ± 2.1	9.4 ± 1.8	36.2	0.003
이야기 복술 정확률(%)	82.7 ± 6.5	59.3 ± 7.1	39.5	0.001
상호작용 참여도(회/세션)	6.5 ± 1.2	5.0 ± 1.1	30	0.002
적극적 감정 점수(1~5점)	4.3 ± 0.4	3.35 ± 0.5	28.4	0.004

데이터 분석 결과는 표 4에 제시되어 있다. 실험군은 평균 독서 지속시간, 이야기 복술 정확률, 상호작용 참여도, 적극적 감정 표현 등 다 항목 지표에서 모두 대조군보다 현저히 높은 수치를 보였다($p < 0.05$). 그 중이 이야기 복술 정확률은 약 23.4% 향상되었고, 상호작용 참여도는 30% 가까이 향상되었으며, 적극적 감정 점수는 28% 이상 향상되었다. 이러한 결과는 데이터 기반 및 인공지능 보조 그림책 디자인 방법이 이야기 완성성, 화면 통일성, 편집 가독성, 상호작용 정확성 등의 측면에서 모두 현저한 우위를 보이며, 아동의 독서 몰입도와 학습 적극성을 효과적으로 강화할 수 있음을 보여준다.

5. 결론

본 연구는 '어떻게 빅데이터와 시각화 기술을 활용하여 아동 그림책 디자인을 최적화할 수 있는가?'라는 핵심 문제를 중심으로 전개되었으며, 아동 그림책에 대한 독서 경험, 시각적 미적 요구, 상호작용에 대한 사용자 요구를 체계적으로 파악하고 분석하여, 그림책 창작과 재 디자인에 있어 실용적이고 실행할 수 있는 의사결정 근거를 제시하는 것을 목표로 하였다. 연구는 먼저 '내용 구성—시각 및 삽화—문자 및 언어—판형 및 레이아웃—상호작용 및 경험—시장 피드백'의 6가지 차원을 기준으로 아동 그림책 키워드 클라우드를 구축하였으며, 이는 공공 도서관 데이터베이스, 전자상거래 플랫폼 사용자 리뷰, 아동 독서 포럼 등에서 수집된 실험 텍스트 데이터를 바탕으로 하였다. 그 후, 텍스트 전처리, 키워드 가중치 계산 및 단단계 시각화를 통해 상호작용이 가능한 키워드 클라우드를 생성하고, 다양한 연령대 사용자들이 가장 중요하게 여기는 그림책 디자인 요소 등을 구조적으로 제시하였다.

이 연구는 그 후, 기존 그림책인 <Teddy의 우산>의 데이터 기반 리 디자인과 주제 기반의 창작 그림책 <뽀빠 혼자 자기>를 시뮬레이션 창작하여, 고빈도 키워드를 활용해 이야기 구조, 플롯 진행, 색상 배합, 삽화 세부 사항, 글꼴 스타일, 레이아웃 비율 및 상호작용 방식을 최적화하고 혁신하였다. 디자인 과정에서는 키워드 클라우드가 창의적 요소를 추출하는 도구로서뿐만 아니라 색조, 이미지 스타일 및 상호작용 영역 배치의 정량적 참조를 제공하였고, 인공지능을 활용한 삽화 생성과 사람과 기계의 협업 수정 과정을 통해 디자인의 효율성 및 일관성을 높였다.

실험 검증 단계에서는 목표 연령층의 아동들을 대상으로 대조군과 실험군을 설정한 독서 상호작용 테스트를 진행한 결과, 데이터 기반 디자인 버전이 이야기의 연계성, 시각적 매력, 글자 가독성, 레이아웃의 편안함, 상호작용 정확성, 독서 몰입도, 감정 긍정성, 이야기 재현을 등 여러 측면에서 유의미한 향상을 보였으며, 이는 해당 방법이 아동 그림책 디자인 분야에서 효과적이고 확장할 가능성을 입증하였다.

그림책 창작자들에게 본 연구는 빅데이터 분석, 키워드 시각화 및 인공지능 보조 생성의 시스템화된 디자인 경로를 제공하였다. 전통적인 방식인 경험과 주관적인 판단에 의존하는 것에 비해, 이 연구는 연령대별 아동들의 내용, 시각, 언어 및 상호작용에 대한 진정한 선호를 정확히 파악하여 이야기와 시각적 표현의 깊은 일치를 이룰 수 있게 하였다. 이 연구는 이론적으로 '데이터—의미—디자인'의 방법론 프레임워크를 구축하고, 실용적으로는 그 실행 가능성을 제시하였으며, 향후 다중 모달 데이터 기반의 스마트 아동 그림책 창작을 위한 실행 가능한 패러다임을 제시하였다.

본 연구는 데이터 수집과 변환 과정에서 일정 부분의 성과를 보여주었으나, 여전히 여러 한계가 내재하여 있다. 첫째, 데이터 샘플의 연령대 분포와 지역 범위는 더욱 확장될 필요가 있다. 둘째, 키워드 추출 과정에서 문맥과 감정적 색조의 인식을 더욱 정교한 의미 분석 모델을 통해 개선해야 한다. 셋째, 키워드 클라우드와 시각적 의미 간의 맵핑 규칙은 표준화된 체계를 구축할 필요가 있다. 향후 연구에서는 아동 행동 추적과 시선 추적 분석을 결합하여 상 황화된 그림책 경험 데이터베이스를 구축하고, 딥러닝을 활용하여 이야기 구조와 시각적 요소의 자동화 생성 및 최적화를 실현하여 아동 그림책이 교육적, 재미, 개별화 적합성에서 더 높은 수준에 도달할 수 있도록 할 것이다.

참고문헌

1. 남미우, '빅데이터 분석 방법을 적용한 인스타그램 뷰티관련 해시태그 게시물 분석에 관한 융합 연구', 한국과학예술융합학회, 2019
2. 서양, 신인식, '소셜 미디어 빅데이터 분석을 활용한 모션그래픽 로고의 시각적 감성 분석', 조형미디어학, 2023
3. 성광숙, 'Z세대 패션에 대한 소셜미디어의 빅데이터 분석', 한국의상디자인학회지, 2020
4. 윤수인, 한정민, 박해신, 김지인, '트렌드 빅데이터 분석 기반의 AI 패션 플랫폼 UX/UI 디자인 개발', 아시아태평양융합연구교류논문지, 2024
5. 장진희, '빅데이터를 활용한 주얼리 마케팅 제안', 한국과학예술융합학회, 2015
6. 경은희, 최정민, 'AI 디자인 툴의 디자인 업무 지원 방향성 제안', Archives of Design Research, 2022
7. 최효승, 손영미, '인공지능과 예술창작 활동의 융복합 사례분석 및 특성 연구', 한국과학예술융합학회, 2017
8. 황윤정, 'BI개발을 위한 생성형 인공지능 디자인 플랫폼 활용에 관한 연구 -'미드저니'의 활용 가능성을 중심으로-', 한국디자인문화학회지, 2023
9. Corey Schimpf, Molly H. Goldstein, 'Large data for design research: An educational technology framework for studying design activity using a big data approach', Frontiers in Manufacturing Technology, 2022
10. Gabrielle A. Strouse, Patricia A. Ganea, 'Parent-Toddler Behavior and Language Differ When Reading Electronic and Print Picture Books', Frontiers in Psychology, 2017
11. HsiaoYuan Hsu, Xiangteng He, Yuxin Peng, Hao Kong, Qing Zhang, 'PosterLayout: A New Benchmark and Approach for Content-Aware Visual-Textual Presentation Layout', 『2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)』, IEEE, 2023
12. Huafeng Quan, Shaobo Li, Changchang Zeng, Hongjing Wei, Jianjun Hu, 'Big Data and AI-Driven Product Design: A Survey', Applied Sciences, 2023
13. Jianan Li, Jimei Yang, Jianming Zhang, Chang Liu, Christina Wang, Tingfa Xu, 'Attribute-conditioned Layout GAN for Automatic Graphic Design', arXiv, 2020
14. Jessica L. Montag, Michael N. Jones, 'The Words Children Hear', Sage Journals, 2015
15. Mahdi Fahmideh, Ghassan Beydoun, 'Big data analytics architecture design—An application in manufacturing systems', Computers & Industrial Engineering, 2019
16. Marzia Mortati, Stefano Magistretti, Cabirio Cautela, Claudio Dell'Era, 'Data in design: How big data and thick data inform design thinking projects', Technovation, 2023
17. Min Zhou, Chenchen Xu, Ye Ma, Tiezheng Ge, Yuning Jiang, Weiwei Xu, 'Composition-aware Graphic Layout GAN for Visual-Textual Presentation Designs', 2022
18. Patricia A. Ganea, Megan Bloom Pickard, Judy S. DeLoache, 'Transfer between Picture Books and the Real World by Very Young Children', Journal of Cognition and Development, 2008
19. Weiwei Cui, Xiaoyu Zhang, Yun Wang, He Huang, Bei Chen, Lei Fang, Haidong Zhang, Jian-Guan Lou, Dongmei Zhang, 'Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements', arXiv, 2019
20. www.cognitivemarketresearch.com