

공공도서관의 잠재 수요 개발에 관한 탐색적 연구*

- 민원게시글과 유튜브 댓글 데이터를 중심으로 -

Exploring Latent Demand for Public Libraries through Civil Complaints and YouTube Comments

윤 명 희 (Myung Hee Yoon)**

강 근 영 (Keun Young Kang)***

목 차

- | | |
|---------------|--------------------|
| 1. 서 론 | 4. 분석 결과 |
| 2. 이론적 배경 | 5. 공공도서관 마케팅 적용 방안 |
| 3. 데이터 수집과 분석 | 6. 결 론 |

초 록

정보환경의 급격한 변화와 다양한 대체재의 등장으로 도서관 마케팅은 생존을 위한 필수 요소가 되었고, 도서관이 시민의 삶에 지속적으로 공존하기 위해서는 서비스 기획의 전 과정에서 시민의 목소리를 반영하는 데이터 기반의 접근이 필요하다. 본 연구는 민원 데이터와 유튜브 댓글이라는 비정형 데이터를 통해 도서관에 대한 잠재 수요 발굴을 외부 마케팅 관점에서 탐색하였다. 이를 위해 2024년 민원 데이터 8,186건과 유튜브 댓글 1,993건을 수집·분석하여 SBERT 기반의 문장 임베딩과 클러스터링 기법을 순차적으로 적용하여, 민원 데이터에서 5개, 유튜브 댓글에서 4개의 군집을 도출하였다. 코사인 유사도 분석을 통해 이들 클러스터를 '정서공감형', '정보탐색형', '정보요구 및 소통형', '제도활용형', '현실참여형'이라는 5가지 세그먼트로 통합하였다. 이들 세그먼트는 표면적 요구를 넘어 정체성 형성, 소속감, 시민 참여 등의 복합적인 동기를 드러내었다. 본 연구는 탐색적 연구로서 정보 요구 기반의 세그먼트를 도출하고, 각 세그먼트별 차별화된 마케팅 접근의 필요성을 제시하였다. 이러한 접근은 시민의 일상적 요구에서 도서관 서비스 확장 기회를 발굴하는데 새로운 방법론적 가능성을 제안하였다는데 의의가 있다.

ABSTRACT

This study explores discovering latent library service demands through unstructured data analysis. We analyzed 8,186 civil complaints and 1,993 YouTube comments from 2024 using sentence-BERT embedding and clustering techniques. Five clusters from complaints and four from comments were integrated into five user segments: emotional empathy, information seeking, information needs and communication, institutional utilization, and civic participation types. These segments revealed complex motivations including identity formation, sense of belonging, and civic engagement beyond simple information needs. This exploratory study demonstrates the potential for identifying latent public library demand through online data analysis, suggesting information needs-based segmentation can enable targeted service strategies beyond traditional demographic approaches.

키워드: 공공도서관, 잠재요구, 텍스트 마이닝, 마케팅 전략

Public Library, Latent Needs, Text-Mining, Marketing Strategy

* 이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2022S1A5C2A03093597).

** 연세대학교 대학도서관발전연구소 연구교수(mhee@yonsei.ac.kr / ISNI 0000 0004 8306 771X) (제1저자)

*** 연세대학교 문헌정보학과 박사후연구원(ky.kang@yonsei.ac.kr / ISNI 0000 0005 0561 2679) (교신저자)

논문접수일자: 2025년 7월 21일 최초심사일자: 2025년 7월 29일 게재확정일자: 2025년 8월 11일

한국문헌정보학회지, 59(3): 203-232, 2025. <http://dx.doi.org/10.4275/KSLIS.2025.59.3.203>

© Copyright © 2025 Korean Society for Library and Information Science

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서론

1.1 연구의 필요성 및 목적

디지털 전환, 인구 구조 변화, 문화 소비 양상의 다변화, 지역 사회의 분권화 등 급변하는 기술·사회·문화·경제적 환경 속에서 공공도서관은 그 존재 이유와 사회적 정당성에 대해 끊임없이 도전받고 있다. 과거의 도서관 직원들은 도서관의 혜택이 명백하고 분명하기 때문에 적극적인 홍보가 없어도 이용자는 도서관을 방문할 것이라고 생각하였다(Kumber, 2004). 그러나 정보환경의 급격한 변화와 도서관 비용의 증감 및 예산 삭감, 이용률 감소 등 다양한 내·외부 환경 변화는 도서관 역할에 대한 재편을 요구해 왔고, 정보시장의 경쟁이 치열해짐에 따라 도서관 마케팅은 생존을 위한 필수 요소가 되었다(Chandratre & Chandratre, 2015).

마케팅은 단순히 광고, 홍보, 판촉 활동만의 영역이 아니라 시민의 요구사항을 효율적으로 식별·예측하고 공급하는 관리 프로세스로서 시민의 삶의 질 향상을 위해 도서관이 제공하는 가치를 극대화하고 이를 효과적으로 전달하는 활동 전반을 포괄한다(Germano & Stretch-Stephenso, 2012). 마케팅의 두 가지 목표는 탁월한 가치를 창출하여 신규 고객을 만들어 내는 것이고, 기존의 고객에게 만족을 제공하여 관계를 강화하는 것이다(Kotler & Armstrong, 2020). 이런 측면에서 공공도서관은 시민의 이용 경험에 기반한 서비스 전략을 넘어서 일상에 잠재된 요구와 태도를 포착하고 이에 대응하는 마케팅 전략 설계가 필요하다.

그동안 도서관에 대한 시민의 요구 분석은 대

출책수, 희망도서 신청수요, 만족도 조사와 같은 설문조사 방식으로 이루어졌으며(최상희, 2021; 심지영, 2021), 이용자 세분화는 인구통계학적 기준이나 이용행태를 통해 분석되었다(신선아, 이명규, 2020; 강미희, 홍현진, 2012). 이런 접근은 도서관을 인지하고 있음에도 이용하지 않는 시민들이나 도서관에 대해 복합적 인식을 가진 시민들의 내면적 동기와 언어를 충분히 설명하기 어렵다는 한계가 있다(Goffin & Lenke, 2004; Bruce, 2010).

비즈니스 영역에서는 표현된 요구 이면의 잠재 요구를 파악하기 위해 SNS와 같은 디지털 상에 남겨진 데이터를 실시간으로 수집·분석하여 마케팅 도구로 활용하고 있다(김형중, 김진화, 2017; Eugenio-Martin & Cazorla-Artiles, 2020). SNS와 같은 소셜 데이터는 매우 많은 사용자들로부터 대량의 데이터를 생성하고 특정 시기 시민들의 즉각적인 반응을 나타내 사회 동향이나 트렌드를 파악할 수 있으며, 사용자의 의도나 문맥에 따라 다양하게 표현되는 비정형화된 데이터로서 구조화된 데이터에서 드러나지 않는 사용자의 욕구와 기대를 포착할 수 있다(이정미, 2013; 최지원, 박승진, 2024).

본 연구는 이러한 문제의식 속에서 공공 민원 데이터와 유튜브 댓글과 같은 비정형 텍스트 데이터를 통해 공공도서관에 대한 잠재요구를 탐색하고자 한다. 민원 데이터와 유튜브 댓글은 기존 조사방법으로 포착하기 어려운 요구와 기대를 자연스럽게 드러내는 디지털 담론 공간으로 도서관 서비스의 새로운 가능성을 발견하고 시민의 삶 속 문제와 연결되는 도서관의 메시지를 설계하는 마케팅 도구로서 의미가 있을 것이다.

1.2 연구방법 및 범위

본 연구는 탐색적 연구로서 외부마케팅 관점에서 온라인상의 데이터를 통해 공공도서관의 잠재수요를 발견하여 마케팅 전략으로 연결 가능성을 모색해 보고자 하였다. 연구방법은 다음과 같다. 첫째, 문헌연구를 통해 도서관 마케팅 이론과 잠재수요 개발 방법을 조사하였다. 둘째, 2024년 민원 데이터 10,182건 중 8,186건을 수집하여 KR-SBERT 문장 임베딩, UMAP 차원축소, K-means 클러스터링을 통해 잠재 고객 세그먼트를 도출하였다. 개별 텍스트의 전체 맥락을 보존하여 기존 토픽모델링으로는 포착하기 어려운 복합적 정보 요구를 분석하였다. 셋째, 두 데이터의 클러스터 간 의미적 유사성을 분석하여 통합 세그먼트를 구성하고 각 세그먼트의 정보요구 특성을 바탕으로 맞춤형 도서관 마케팅 전략을 제안하였다. 연구범위는 2024년 데이터로 제한하며, 분석 대상은 민원과 유튜브라는 특정 플랫폼의 텍스트 데이터로 한정하였다. 이러한 접근 방식은 설문조사 등 기존의 수요조사 방식과 차별화되는 잠재수요 발굴의 방법론적 토대를 제시한다는 점에서 의의가 있다.

2. 이론적 배경

2.1 외부지향 마케팅(Outside-in Marketing)의 필요성

최근 비즈니스 영역에서는 내부 효율성이나 자원 중심의 Inside-out 접근과 달리 외부 이해

관계자의 신호에 민감하게 반응하고 전략에 반영하는 능력을 강조하는 Outside-in 마케팅이 강조되고 있다(Mu et al., 2018; Musarra & Morgan, 2020; Rust, 2020). Outside-in 마케팅은 조직의 전략이 내부에서 출발하는 것이 아니라 외부 고객의 요구에서 출발하는 것이 경영 수익에 유리하다는 관점이다. 이는 단순한 시장지향 행동의 집합이 아니라, 조직문화와 활동을 포함한 복합적인 개념이다(Musarra & Morgan, 2020). 외부 지향 마케팅을 성공적으로 실행하기 위해서는 전통적 시장 조사에서 벗어나 빅데이터, 적응형 개인화 시스템(adaptive personalization), 소셜미디어 마이닝을 활용한 실시간 고객 모니터링, 브랜드 관리자 중심에서 고객관리자 중심으로 조직 재편, 단기 매출 중심의 성과지표에서 고객생애가치, 고객자산 등 장기적 고객 중심지표로의 전환이 필요하다(Rust, 2020). 외부 지향적 마케팅 역량은 시장 감지(Market Sensing), 고객관여(Customer Engaging), 파트너 연결(Partner Linking)의 세가지 핵심요소로 구성된다(Mu et al., 2018). 이러한 외부 지향적 마케팅은 막대한 양의 고객 정보 데이터를 분석하여 시장의 변화를 파악해야 하고, 이를 위한 최신의 지식 확보가 필수적이다. 정확한 지식과 업데이트된 정보를 갖추지 못하면 효과적 마케팅 의사결정이 어렵기 때문에 조직은 정보전략을 최신화하여 마케팅 활동을 지원해야 한다(Varadarajan, 2020; Styf, 2020).

이같은 외부 지향 마케팅은 도서관에 시사하는 바가 크다. 경쟁적이고 역동적인 외부 환경에서 도서관이 경쟁력을 확보하려면 전통적 내부 중심 전략에서 벗어나 외부 고객 중심의 전

략적 관점을 채택할 필요가 있다. 외부지향적 관점은 조직의 자원과 역량을 외부 고객의 요구를 파악하고 충족시키기 위해 전략적으로 활용하는 것을 의미하고, 이러한 역량을 강화함으로써 조직이 지속적인 경쟁우위를 확보할 수 있음을 시사한다.

이 관점에서 도서관에서 필요한 시장감지 역량은 외부 환경의 변화와 새로운 기회를 예측하고 탐지하는 능력으로 환경변화에 맞춰 조직 내부의 역량을 업데이트 하는데 필수적이다. 기존의 대출 책 수, 이용자 수 이외의 다양한 외부 데이터를 활용하여 기존 서비스에서 충족되지 못한 새로운 서비스 기회를 포착할 수 있다. 고객관계 역량은 고객과 지속적으로 상호 작용하고 관계를 구축하는 능력으로 고객의 잠재된 요구를 탐색하는 데 중요한 역할을 한다. 기존의 정형화된 설문조사나 포커스그룹 이상의 고객관찰과 공감적 접근법을 활용하여 고객 스스로 인지하지 못하는 숨겨진 요구를 발견할 수 있어야 한다. 파트너 연결은 도서관 외부 조직과의 협력을 통해 상호보완적인 자원을 조정하고 활용하여 고객가치를 창출하는 능력이다. 외부 파트너와의 연계를 통해 새로운 서비스 아이디어를 얻고, 외부 자원을 활용하여 보다 혁신적이고 고객맞춤형 서비스를 개발할 수 있을 것이다(Mu et al., 2018).

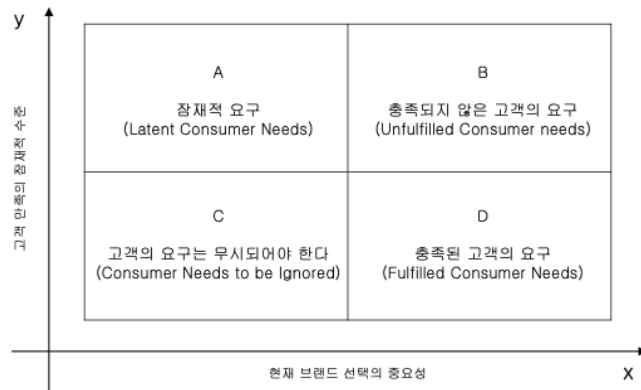
이러한 역량은 현재 도서관을 이용하고 있으나, 만족하지 못한 이용자의 불만족 원인과, 도서관을 이용하지 않는 비이용자의 장애요인, 이용이 중단된 이용자의 비이용 원인 등 밖으로 표현되지 않았지만 도서관과 연결된 잠재된 요구를 발굴할 수 있는 기본역량이다.

2.2 잠재요구의 개념

외부 지향 마케팅은 고객의 요구를 전략 개발의 출발로 삼고 있다. 마케팅 역량이 탁월한 조직은 고객의 요구(Needs)를 배우고 이해하기 위해 전력을 다한다(Kotler & Armstrong, 2020). 어떤 조직이든 경쟁에서 살아남기 위해서는 고객의 요구에 민감해야 하고 이들의 요구를 파악할 수 있는 다양하고 방대한 양의 고객 데이터를 축적해야 한다. 고객의 데이터는 조직의 핵심 자산이다(Rosário & Dias, 2023).

그러나 고객의 요구는 고객조차 명확하게 표현하지 못할 때가 많다(Goffin & Lemke, 2004; Bruce, 2010). 고객이 스스로 인지하고 표현할 수 있는 요구가 명시적(explicit) 요구라면, 고객조차 명확하게 인식하지 못하거나 표현하지 못하는 무의식적 수준의 요구를 잠재요구(Latent Needs)라고 한다(Das, 2012; Bruce, 2010). 조직은 고객의 잠재된 요구와 기대를 명확히 규명함으로써 조직 성장과 지속 가능성을 이끌어낼 수 있다. 잠재요구는 고객이 원하는 카테고리의 새로운 동인을 의미하며, 혁신과 새로운 서비스 발견의 원천이다(Das, 2012). 〈그림 1〉은 비즈니스 영역에서 고객이 특정 브랜드 선택 시 나타나는 의식적인 요구를 나타낸 것으로, 도서관에 대한 고객의 요구에도 적용될 수 있다.

x축은 고객의 특정 속성에 대한 의식적 요구를 나타내고, y축은 해당 요구가 충족되었을 때 미래에 달성될 고객 만족의 잠재적 수준을 나타낸다. A영역은 현재 어떤 도서관도 이 요구를 충족하지 못하기 때문에 고객은 도서관을 선택하지 않는다. 하지만 미래에 도서관이 이



〈그림 1〉 Framework for different type of consumer needs(Das, 2012)

요구를 충족하게 된다면 고객은 도서관을 찾게 될 것이다. 이 영역은 잠재고객을 나타내며, 미래에 고객을 만족시킬 수 있는 욕구를 충족시켜 장기적인 잠재력을 제공할 수 있다. B영역은 충족되지 않는 고객의 요구로서 현재 도서관을 이용하고 있으나, 욕구가 충족되지 않아서 기존 서비스를 보완하거나 새로운 서비스를 개발하여 이러한 요구를 충족시킬 때 만족도를 높일 수 있는 잠재적 영역이다. D영역은 충족된 요구로서 현재 도서관에 충분히 만족한 고객으로서 관계를 강화하여 고객옹호와 지지를 획득해야 할 영역이다. C영역은 무시해도 되는 요구로서 현재 중요하지 않고 미래에도 만족도가 낮을 것으로 예상되며 현재 또는 미래의 성장 잠재력 측면에서 큰 이점을 제공하지 않는 영역이다. 다만 이 영역에는 도서관과 연결되어 있으나 전통적인 도서관의 정체성에 부합하지 않는다는 이유로 외면받아 온 요구도 포함될 수 있으며, 이러한 요구들은 넓은 의미에서 본 연구의 잠재요구 범주에 포함될 수 있다. 따라서 A, B, C영역의 잠재요구를 충족시킨다면 도서관은 신규 수요를 창출함과 동시에 기존

고객의 만족도를 높일 수 있을 것이고 D영역에서는 만족한 요구를 끌어올려 고객옹호와 지지를 이끌 것이다.

본 연구는 이러한 이론적 프레임워크를 바탕으로 A영역의 잠재요구 발굴에 초점을 맞추어 도서관 외부의 다양한 온라인데이터를 활용한 잠재수요 개발을 제안하고자 한다.

2.3 잠재요구 측정 방법

잠재요구를 파악하는 데 가장 어려운 점은 현재 도서관 영역 밖에 있는 사람들의 일상적 요구 속에서 도서관의 필요성을 이끌어 내는 방법을 찾는 것이다. Goffin과 Lemke(2004)은 기존의 설문조사와 포커스 그룹 인터뷰 등은 고객이 인지하지 못하거나 표현하지 못하는 은연중의 숨겨진 요구(Hidden consumer needs)를 발견하는 데 한계가 있다고 했다. 성공적인 혁신 서비스 개발을 위해서는 표면적으로 드러난 요구를 넘어 더 깊은 고객의 요구를 포착하는 새로운 접근법이 필요하다. Bruce(2010)은 숨겨진 소비자 요구를 정확히 발견하는 것은 혁

신의 성공 가능성을 높이고 경쟁력 향상에 기여하기 때문에 효과적인 방법을 찾는 것이 중요하다. 언급하면서 관련 연구문헌 29편을 분석하여 잠재 요구 식별에 사용된 주요 연구방법을 소개하였다. Goffin과 Lemke(2004)이 심리학, 인류학 등에서 개발된 HNA(Hidden Needs Analysis) 도구와 Bruce(2010) 연구에서 제시한 잠재요구 분석 방법을 정리하면 <표 1>과 같다.

Bruce(2010)는 잠재수요 발굴을 위한 4가지 접근방법을 제시했는데, 이는 도서관에도 적용될 수 있다. 첫째, 현 제품이나 서비스의 불편한 점을 관찰 기록하여 숨겨진 요구를 추론하는 문제지향(Problem-oriented) 접근이다. 이는 HNA에서 강조하는 관찰과 공감적 디자인 등의 방법을 활용한다. 둘째, 고객의 개인적 사회문화적 가치관을 이해해 숨겨진 요구를 파악하는 가치지향(Values-oriented) 접근이다. 이는 문화적 맥락을 파악하기 위해 인류학적 조사나 계층화(Laddering) 인터뷰가 사용될 수 있다. 셋째, 고객을 설계 과정에 적극 참여시키는 참여지향

(involvement-oriented) 접근이다. 이는 프로브 앤 런(probe and learn)이나 미니 콘셉트 테스트 같은 도구를 통해 잠재 수요를 도출한다. 넷째, 새로운 기술을 제시해 사용자에게 현재의 서비스에서 느끼지 못했던 공백을 깨닫게 하는 기술지향(technology-oriented) 접근이다.

한편, 기술의 진보와 소셜 플랫폼 활성화에 따라 디지털 공간에 고객들이 자발적으로 생성한 데이터가 마케팅에서 중요하게 활용되고 있다(Islam & Sheikh, 2024; Odung & Agungi, 2023). 소셜미디어는 고객의 무의식적 인식 구조를 반영하는 주요한 도구로서 마케팅에서 고객의 잠재된 기대와 요구를 파악하는데 유용하다(김형중, 김진화, 2017). Rosário와 Dias(2023)은 기술의 진보와 함께 기업이나 조직은 데이터 기반 마케팅을 통해 고객의 행동 분석, 개인화, 고객만족, 전환율 개선 등에서 긍정적 효과를 발휘하고 있다고 했다. 일찍이 전채남과 서일원(2013)은 빅데이터 분석 기술을 활용하여 기존에 파악되지 못한 잠재 수요 기업을 발굴하

<표 1> 잠재요구를 발견하는 방법

| 연구자 | 분석방법 | 내용 |
|--------------------------|--|---|
| Goffin & Lemke (2004) | 레퍼토리 격자 분석법 (Repertory Grid Analysis) | 심리학에서 개발된 기법으로 간접적 질문을 통해 고객의 인지·비인지적 요구 탐색 |
| | 공감적 디자인(Empathic Design) | 인류학적 관찰 기법 활용해 사용자 행동, 몸짓, 환경 맥락에서 숨은 요구 포착 |
| | 맥락적 면담(Contextual Interviews) | 실제환경에서의 반구조화된 인터뷰로 현실적 사용 문제와 기대 도출 |
| | 선도 사용자 그룹(Lead User Groups) | 특수 사용 조건에 있는 리드 유저의 요구와 개선 아이디어를 통해 일반 시장 혁신 유도 |
| Bruce (2010) | 민족지학(Ethnography) | 실제 고객 일상 관찰을 통해 숨겨진 문제와 문화적 가치 파악 |
| | 사다리법(Laddering Interview) | 제품 속성의 중요성에 대한 심층 자문을 통해 소비자의 개인적 가치와 무의식적 욕구도출 |
| | 고객참여(Customer Involvement) | 고객과 지속적 상호작용으로 요구 직접 발굴 |
| | 미니 콘셉트 테스트(Mini Concept Test) | 새로운 제품 기능 컨셉을 고객에게 제시해 요구 자극 |

고 이 데이터를 기반으로 기술 마케팅 전략 수립을 강조하였고, 이정미(2013)는 빅데이터는 현대인의 일상적 기대와 욕구를 드러내는 총체적 본산이며 현실 정보세계를 사회의 관점에서 볼 수 있는 유용한 도구라고 하여, 이를 활용한 데이터 기반 도서관 정보서비스의 필요성을 강조하였다. 이러한 배경에서 텍스트 데이터를 통한 잠재 요구 분석을 위해 활용되는 주요 기법들을 정리하면 <표 2>와 같다.

도서관은 이와 같은 질적·양적 다양한 방법을 통해 단순히 이용자 요구에 응답하는 수준을 넘어 시민의 일상 속에 잠재된 결핍을 감지하여 서비스로 제안하는 역동적 기관으로 진화할 필요가 있다. 이를 위해서는 외부지향적 관점에서 외부의 신호를 민감하게 감지하고 이를 내부 전략으로 연계하여 서비스 기획 실행으로 만들어 지속가능한 수요 창출형 마케팅 전략이 필요하며, 이는 외부데이터를 분석하고 해석·활용할

수 있는 역량, 조직간 협력 및 서비스 연계 체계 구축이 병행되어야 한다(Mu et al., 2018).

2.4 선행연구

도서관 마케팅 전략은 전통적으로 서비스 홍보나 프로모션 중심으로 접근되어 왔다. 초기 연구들은 도서관 이용자 요구를 파악하기 위한 마케팅 전략에 주목하였다. 한계문(2003)은 인터넷 기반 마케팅 전략 수립을 위해 시장세분화 기법의 도입 가능성을 제시하였으며, 이용재(2012)는 뉴욕공공도서관 사례를 분석하여 체계적 마케팅 전략이 공공도서관의 가치를 강화하는 데 효과적임을 보여주었다. 이성신(2014)은 전략기획 세분화, 타겟설정, 포지셔닝의 마케팅 전략 모델을 도서관 환경에 적용하는 모형을 제시하여 도서관 운영에 마케팅 체계화 필요성을 강조하였다. 김미설과 장우권(2015)은 마

<표 2> 텍스트마이닝 관련 분석기법

| 구분 | 분석기법 | 특징 및 적용 목적 |
|-------------|----------|--|
| 전통적 기법 | 토픽모델링 | 문서 집합에서 잠재 주제 추출 대량 텍스트의 주제 분류 및 트렌드 파악 |
| | 키워드 추출 | 핵심 키워드 추출 및 빈도 분석 주요 이슈 발굴, 관심사 파악 |
| 의미 분석 기법 | word2vec | 단어의 벡터 표현 학습 의미적 유사성 기반 분석 |
| | Doc2Vec | 문서 전체의 벡터 표현 학습 문서 간 유사성 분석 |
| | BERT | 문맥을 고려한 양방향 언어 모델 복합적 의미, 맥락적 정보요구 포착 |
| | SBERT | 문장 단위 의미 표현 이질적 텍스트 통합 분석, 세그먼트 도출 |
| 감정/의도 분석 기법 | 감성 분석 | 텍스트의 긍정/부정 감성 분류 고객 만족도, 브랜드 인식 분석 |
| | 오피니언 마이닝 | 텍스트에서 의견과 주관적 정보 추출 사용자의 의견 및 태도 분석 |

케팅 논문 분석을 통해 대부분의 연구가 이벤트 프로그램, 홍보물 제작, 이용자 참여 유도 등 프로모션 중심의 단편적 접근으로 이루어졌다고 밝혔다.

한편, 디지털 환경에서 발생하는 비정형 데이터의 가치는 꾸준히 강조되어 왔다. 이정미(2013)는 도서관에서의 빅데이터 활용 가능성을 조망하며 방대한 외부 데이터를 통한 이용자 행동 예측 및 수요 분석의 가능성을 제시하였다. 표순희 외(2015)는 공공도서관 분야에 적용가능한 빅데이터 활용가치를 규명하고, 수요자 요구 분석을 기반으로 빅데이터 서비스 모형을 개발하여 장기적으로 빅데이터 기반 서비스가 도서관 업무 효율성, 맞춤형 서비스 제공, 지역사회와의 연계성 강화에 기여할 것으로 파악하였다. 2020년 이후 발표되는 연구에서는 소셜 데이터 기반 이용자 요구 분석 연구들이 등장했다. Zou et al.(2020)은 미국 전역의 10개 공공도서관 트위터 계정에서 10,000개 이상의 트윗을 수집하여 토픽 모델링과 감성분석이라는 텍스트 마이닝 기법과 트위터 계정 이용자 설문조사를 병행하여 도서관이 어떻게 효과적으로 이용자 참여와 상호작용 확대 전략을 개발할 수 있는지를 실증적으로 탐구했다. 박태연과 오효정(2021)은 코로나19 시기 OpenAPI를 활용하여 자체 개발한 크롤러를 사용하여 496,741건의 도서관 관련 트위터 데이터를 수집하여 시계열 기반 주요 이슈 탐지 및 트래킹(TDT: Topic Detection and Tracking) 등의 텍스트 마이닝 분석을 실시하였다. 도서관 운영 주요 이슈에 따라 기간을 구분하고 기간별 주요 용어들의 출현 빈도의 변화를 측정하여 이슈별 용어 가중치를 비교·분석하여 잠재 이용자의 요

구를 4가지 이슈로 구분하고 각 이슈별 시사점을 도출하였다. 최지원과 곽승진(2024)은 도서관 SNS에 나타난 이용후기 데이터와 온라인 뉴스 기사를 텍스트마이닝 기법으로 분석하여 이용자 인식과 평가를 탐색하였다. 이를 통해 정책적 사회적 흐름 속에서 지역사회 수요에 기반한 도서관 운영 전략 수립 방안을 모색하였다. 이러한 접근은 단순히 통계 수치를 넘어, 정서적 반응과 사회적 의미를 함께 고려한 이용자 요구 파악이라는 점에서 수요기반 전략 수립의 실마리를 제공하였다. 그러나 이 연구들은 현재 도서관과 연결된 고객의 요구에 기반한 데이터 수집·분석에 초점을 맞추고 있어, 도서관 서비스 밖에 존재하는 고객들의 요구를 발견하고 분석하여 새로운 수요를 창출하는 마케팅 전략으로 연결하는 데까지 미치지 못하고 있다.

반면, 비즈니스 영역에서는 온라인 데이터 기반 분석을 마케팅 전략과 연결한 연구들이 다수 존재하였다. 전채남과 서일원(2013)은 산업 및 기술마케팅에서 텍스트마이닝, 시맨틱네트워크 분석 등 빅데이터 분석기법을 활용하여 특정 기술과 연관된 잠재 수요 기업을 발굴하는 방법론을 제시함으로써 데이터 기반 잠재고객 발굴의 실무적 기여를 강조하였다. Satish와 Kavya(2016)은 Hadoop을 활용하여 E-Commerce 트렌드를 분석하고 전자상거래 사이트의 로그 데이터(상품조회, 구매 클릭 등)를 수집하여 고객의 구매 패턴, 선호도, 클러스터를 식별함으로써 기업의 의사결정과 마케팅 전략수립에 연결하였다. 김형중과 김진화(2017)는 소비자 가치 인식을 중심으로 트위터 빅데이터를 분석하고 감성분석 및 토픽모델링을 통해 긍정/부정 가치 유형을 도출하였다. 이는 단순한 요구의

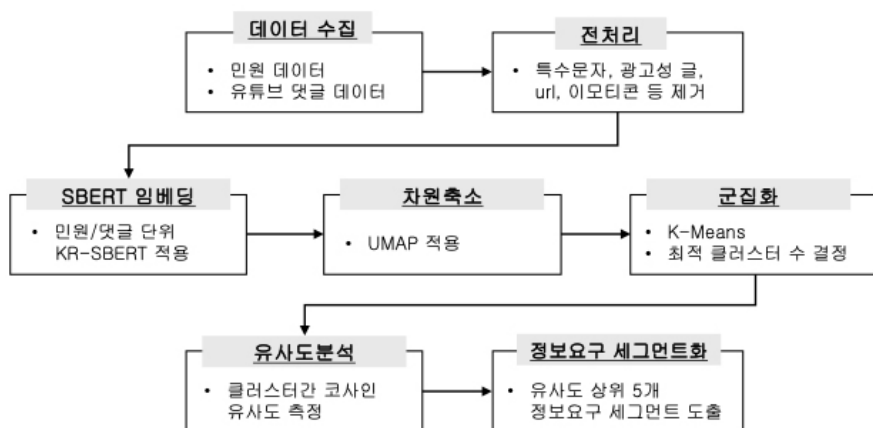
나열을 넘어서, 이용자의 감성 및 태도에 기반한 수요 우선순위 설정이 가능함을 보여준다. Eugenio-Martin과 Cazorla-Artiles(2020)은 중국의 국내 관광 수요에 내재된 잠재적 수요를 파악하기 위해 점유율 방법(Shared Method)이라는 새로운 정량적 측정 도구를 제안하고 그 유효성을 검증하였다. 이는 기존의 관광 수요 예측이 과거의 실현된 수요에 의존했지만, 이 연구는 정책적, 사회경제적, 심리적 요인 등으로 실현되지 못한 수요를 통계적으로 추정했다는 데 의의가 있다. 신호철과 허철무(2022)는 소비자 리뷰 데이터와 머신러닝 기반 분석을 통해 소비자 유형을 분류하고, 각 군집에 따른 특성을 시각화하여 마케팅 전략 수립에 필요한 인사이트를 제공하였다. 주가혜 외(2023)는 장소 브랜딩 전략 수립을 위해 소셜 미디어 플랫폼(Ctrip)의 대중 코멘트를 수집하고, 파이썬 기반 키워드 추출 및 군집 분석을 통해 장소에 대한 인식 유형을 분류했다. 이들은 마케팅 전략을 ① 자원 개발, ② 이미지 포지셔닝, ③ 가치 창출, ④ 전파 및 확산이라는 4단계로

구분하고, 각각의 단계에서 요구를 반영한 전략 요소(예: 고유 자원형, 체험형, 서사적 공감형)를 도출함으로써, 이용자 감성 기반의 마케팅 전략 수립 가능성을 보여주었다.

본 연구에서는 비즈니스 영역의 관련 연구를 참고하여 도서관의 잠재 수요를 발굴하기 위해 온라인 상의 민원게시글과 유튜브 댓글 데이터를 수집 분석하고 이를 유형화하여 마케팅 전략 설계의 토대를 제공하고자 한다.

3. 데이터 수집과 분석

본 연구에서는 온라인상의 데이터를 활용하여 도서관의 잠재 요구를 발견하고 이를 세그먼트화하여 마케팅 전략화로 이끌기 위한 방법으로 데이터 기반의 프레임워크를 제안하고자 한다. 이에 성격이 다른 두 가지 유형의 자연어 데이터를 대상으로 연구를 수행하였다. 이 연구에서 수행한 데이터 수집과 분석 방법은 〈그림 2〉와 같다.



〈그림 2〉 데이터 수집 및 분석방법

3.1 데이터 수집 및 전처리

본 연구는 다양한 온라인 데이터 중에서도 구체적 문제 해결을 요구하는 민원 데이터(이고은, 김기영, 2014)와 콘텐츠 기반의 자발적 반응인 유튜브 댓글을 수집하였다. 이러한 데이터 조합을 선택한 이유는 문제 인식에서 해결 기대까지의 전 스펙트럼을 포괄할 수 있고 공적 요구와 사적 감정을 동시에 파악할 수 있기 때문이다. 민원 데이터는 명확한 개선 요구와 정책적 실현 가능성을 담고 있어 구체적 서비스 방향을 제시하며(이고은, 김기영, 2014), 유튜브 댓글은 기존 조사방법으로 포착하기 어려운 무의식적 인식과 감정적 반응을 맥락과 함께 자연스럽게 드러내는 자발적 표현 데이터로서 의미가 있다고 판단하였기 때문이다(안성우, 2025; 김승재, 조규락, 2025; 최믿음 외, 2024).

민원 데이터는 한국지능정보사회진흥원이 운영하는 AI 관련 데이터 통합 플랫폼인 AI hub에서 제공하는 '공공 민원 상담 LLM 사전학습 및 Instruction Tuning 데이터'를 활용하였다. 2024년에 구축된 총 10,182건의 원천 데이터 중 중앙행정기관과 지방행정기관으로 등록된 민원 8,186건을 모델링에 사용하였다. 민원 데이터의 경우 주로 공적인 문제 해결과 정보 요청을 목적으로 하며, 상대적으로 길고 구조화된 텍스트로 구성되어 있다. 전문적인 용어도 자주 사용되며 구체적인 문제와 요구 사항이 명확하게 표현되는 특징이 있다. 민원 데이터는 표현된 수요라고 할 수 있지만, 일상 속의 비표현된 관심사를 포착하기 위해, 공공 이슈와 직접적인 관련이 없는 일반 생활 주제를

중심으로 유튜브 댓글 데이터를 수집하였다. 유튜브 댓글 데이터는 검색 키워드를 선정하여 동영상 검색 후 각 키워드별 동영상 20개에 댓글 10개씩 수집하였다. 키워드 선정의 이론적 정당성과 실증적 관심도를 확보하기 위해 통계청과 정부조사에서 활용중인 분류체계(예: 주거/생활, 건강/웰빙, 여가/문화, 교육, 경제/소비 등)를 참고하여 다섯 가지 영역(주거 및 공간, 생활습관 및 루틴, 건강 및 자기관리, 소비생활, 감성 및 라이프스타일)을 설정하였다. 그리고 각 카테고리별로 구글 트렌드의 2024년도 2024년 1월부터 12월까지의 검색량 데이터를 바탕으로 키워드를 선정하여 연간 검색량이 지속적으로 확인된 대표 키워드를 반영하였다. 본 연구에서 활용한 키워드는 다음과 같다:

```
keywords = [
    "원룸", "브이로그", "생활 꿀팁", "동네", "혼자 여행", "하루 루틴", "루틴", "헬스", "자기계발", "스마트폰 꿀팁", "AI 사용 후기", "AI 꿀팁", "절약", "가성비 추천템", "가성비", "추천", "감성", "심리상담", "명상", "마음챙김", "템플릿", "생산성 앱", "자취 레시피", "생활비 관리", "북튜버", "독서 브이로그", "미니멀라이프", "감성 브이로그", "챌린지"
]
```

키워드 검색 후 영상 url을 수집하고 중복 영상을 제거하였다. 495건의 영상에서 수집이 가능한 2,244개의 댓글을 수집하였으며, 이 중 불필요한 요인(광고성 댓글, 타임라인 게시용 등) 삭제 후, 1,993건의 댓글이 모델링 대상이 되었다.

민원 데이터가 사회 전반의 보편적인 관심사를 포괄적으로 조망하기 위한 기준으로 활용이 가능한 반면, 유튜브 댓글 데이터는 개인적인 의견이나 감정 표현을 주목적으로 하며, 짧고 비격식적인 텍스트로 구성되어 있다. 이모티콘이나 축약어, 구어체 표현이 주로 사용되며 즉흥적이고 주관적인 반응이 주를 이룬다. 이와 같이 성격이 상이한 두 데이터 셋을 함께 분석하며 다양한 상황과 맥락에서 표출되는 잠재요구를 포괄적으로 파악하고자 하였다.

전처리된 특수문자, url, 광고성 글, 이모티콘 등 분석에 불필요한 요소와 글을 분석에서 제외하는 과정이었다. 이후 각각의 글이 담고 있는 전체적인 맥락과 의미를 온전하게 보존하면서 분석하기 위하여, 민원 게시글과 유튜브 댓글을 개별 분석의 단위로 설정하기 위한 처리 과정을 거쳤다. 민원의 경우 문제 상황과 요구 상황, 배경 설명 등이 유기적으로 연결되어 있는 경우가 많아, 민원 한 건을 단위로 분석하는 것이 적합하다고 보았으며, 유튜브 댓글의 경우, 문장 단위로 분리하면 그 의미가 파편화될 수 있어 댓글 한 건을 하나의 단위로 분석하여 작성자의 의도와 의미를 보다 명확하게 파악하고자 하였다.

3.2 텍스트 임베딩

전처리된 텍스트 데이터는 SBERT(Sentence-BERT)를 활용하여 고차원 벡터공간에 임베딩하였다. SBERT는 BERT의 변형 모델로, 문장 수준의 의미론적 유사성을 효과적으로 포착할 수 있는 특징을 가지고 있다(Reimers & Gurevych, 2019). 전통적인 텍스트 분석 방법

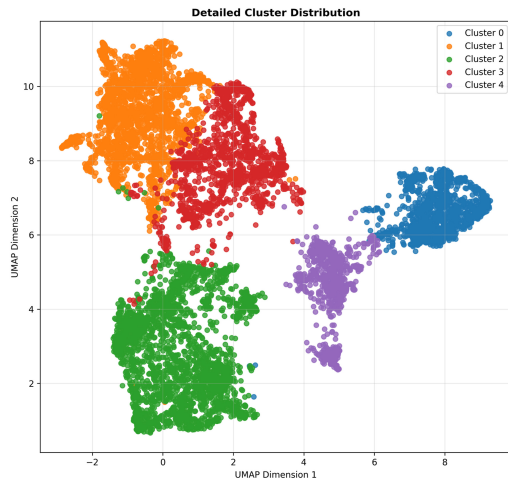
인 토픽모델링과 달리, SBERT는 단어의 맥락과 의미적 관계를 고려하여 문장 전체의 의미를 포착할 수 있어 특히 짧은 텍스트 분석에 더 적합하다(Reimers & Gurevych, 2019). 그리고 사전훈련된 언어 모델을 기반으로 하기 때문에 LDA와 같은 통계 기반의 모델보다 적은 데이터로도 의미있는 결과를 도출할 수 있는 장점이 있다. 본 연구에서는 한국어 기반의 KR-SBERT(Park & Shin, 2021) 모델을 사용하여 민원 게시글과 유튜브 댓글 데이터를 각각 768차원의 벡터로 변환하였고 두 데이터 유형의 언어적 특성 차이를 반영하기 위하여 도메인별로 독립적인 임베딩을 수행하였다.

3.3 차원 축소

다음으로 SBERT를 통해 생성된 고차원 벡터는 시각화와 클러스터링 분석을 위해 UMAP(Uniform Manifold Approximation and Projection)을 통한 차원 축소 과정을 거쳤다. UMAP은 t-SNE 방식과 유사하지만 계산 효율성이 높고 데이터의 구조를 더 잘 보존한다(McInnes et al., 2018). 특히 UMAP은 데이터의 국소적 구조 뿐만 아니라 전체적인 관계도 보존하기 때문에, 서로 다른 문서 간의 의미적 관계를 보다 정확하게 시각화하는데 효과적이다. 본 연구에서는 $n_neighbors=15$, $min_dist=0.1$, $n_components=2$ 으로 파라미터를 설정하였고 768차원 벡터를 2차원으로 축소하였다.

3.4 군집화

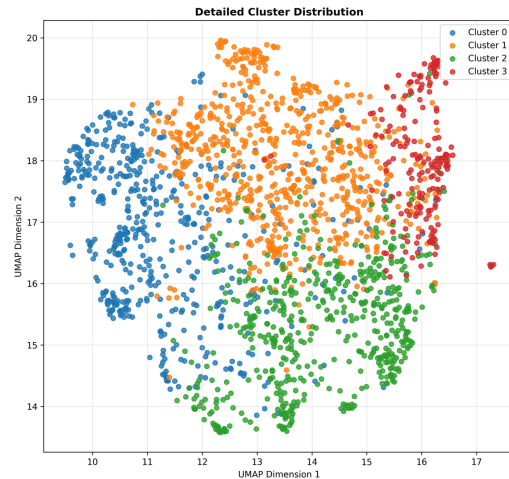
마지막으로 차원이 축소된 데이터에 K-means



〈그림 3〉 민원 클러스터 분포 시각화

클러스터링 알고리즘을 적용하여 의미적으로 유사한 민원과 댓글을 그룹화하였다. 최적의 클러스터 수(k)를 결정하기 위해 k 값 조정 실험 ($k=2\sim 12$)과 함께 실루엣 스코어를 참고로 사용하였으며, 그 결과 민원 데이터는 $k=5$, 유튜브 댓글은 $k=4$ 로 설정하여 클러스터를 분석하였다.

본 연구에서 전통적인 토픽모델링 대신 위의 조합을 선택한 주요 이유는 첫째, 민원 게시글과 유튜브 댓글의 경우 길이가 짧고 비정형적인 특성이 있어, 단어 빈도에 의존하는 토픽모델링의 성능이 제한적일 수 있다. 둘째, 토픽모델링은 단어의 의미적 유사성보다는 동시 출현 확률에 기반하기 때문에 문맥적 의미를 파악하기 어렵다. 셋째, SBERT는 사전학습된 언어 모델의 강점을 활용하여 적은 양의 데이터에서도 의미적으로 풍부한 표현을 추출할 수 있다. 넷째, 본 연구의 목적인 잠재 요구 세그먼트를 파악하기 위해서는 단순한 주제 분류를 넘어 텍스트에 내재된 심층적인 맥락과 의미적



〈그림 4〉 유튜브 클러스터 분포 시각화

패턴을 이해할 필요가 있으며, 이러한 목적에 SBERT 기반 임베딩 방식이 적합하다 판단하였다.

이때 민원 데이터와 유튜브 댓글이라는 이질적 데이터를 통합 분석하기 위해 다음과 같은 방법론적 고려사항을 적용하였다. 첫째, 전처리 과정에서 각 데이터 유형의 고유 특성을 보존하면서 분석 가능한 형태로 정규화하였다. 둘째, 각 데이터 유형별로 독립적인 클러스터링을 수행한 후, 클러스터 중심점 간 코사인 유사도를 통해 의미적 연관성을 측정하였다. 이후 높은 유사도를 보인 클러스터 쌍은 구성 텍스트에 대한 정성적 검토를 병행하였다.

이러한 방법론적 접근은 비정형적 텍스트 데이터의 구조를 효과적으로 파악하고 잠재적 정보 요구 패턴을 발견하기 유용할 수 있다. 특히 성격이 다른 두 자연어 데이터를 통합적으로 분석하여 도서관 다양한 잠재 요구 세그먼트에 대한 데이터 기반의 이해를 제공할 수 있다.

4. 분석 결과

4.1 클러스터 분석 결과

본 연구에서는 민원 데이터와 유튜브 댓글이라는 이질적 데이터를 통합적으로 분석하기 위해 3장에서 설명한 바와 같이 도메인별로 독립적 임베딩을 수행하여 언어적 특성 차이를 반영하였고, 클러스터 중심점 간 코사인 유사도를 통해 의미적 연관성을 측정하였다. 또한 결과 해석 시 민원의 문제 해결 지향적 맥락과 댓글의 감정 표현 중심 맥락의 차이를 고려하여, 동일 세그먼트 내에서도 표현 방식에 따른 의미 차이가 있을 수 있음을 유의하였다.

민원 데이터(8,186건)와 유튜브 댓글 데이터(1,993건)에 대해 각각 KR-SBERT 기반의 문장 임베딩을 수행한 후, 차원축소와 클러스터링을 적용하여 잠재적 정보 요구를 파악하고자 하였다. 그 결과, 민원 데이터에서는 5개의 클러스터, 유튜브 댓글 데이터에서는 4개 클러스터가 도출되었다. <표 3>과 <표 4>는 각 데이터셋의 클러스터 분포를 나타낸 것이다.

각 클러스터를 살펴보면, 클러스터 0은 평균 121자로 상대적으로 짧은 민원을 포함하며 간결한 민원 패턴을 보여준다. 이는 단순 문의나 요청 사항을 다루는 세그먼트로, 근로조건과 노동 관련 민원이 중심을 이루었다. 클러스터 1은 평균 408자로 상대적으로 가장 긴 민원 길이를 보이며, 민원 복잡성이 다양함을 보여준다. 복잡한 사안이나 상세한 설명이 필요한 민원이 주를 이루며 교통과 도시환경 관련 민원이었다. 클러스터 2는 가장 큰 규모의 클러스터로 평균 394자의 중간 길이를 보인다. 일반적인 민원 처

리 패턴과 관련된 세그먼트로, 부동산 및 건축 법규 관련 민원으로 구성되었다. 클러스터 3은 평균 323자이나 일부 매우 긴 민원이 평균을 끌어올렸으며 대다수는 186자(중앙값) 민원이었다. 단순한 문의와 복잡한 사안이 동일한 의미적 클러스터로 분류되었다. 실제 클러스터에는 행정서비스 및 지원제도 관련 민원이 포함되었다. 클러스터 4는 평균 173자 문자 수를 가지며, 비교적 짧은 길이의 민원이었다. 창업 및 경영 지원 관련 민원으로 구성되었다. 민원 클러스터별 대표적인 예시는 다음과 같다:

클러스터 0 (근로와 노동 관련):

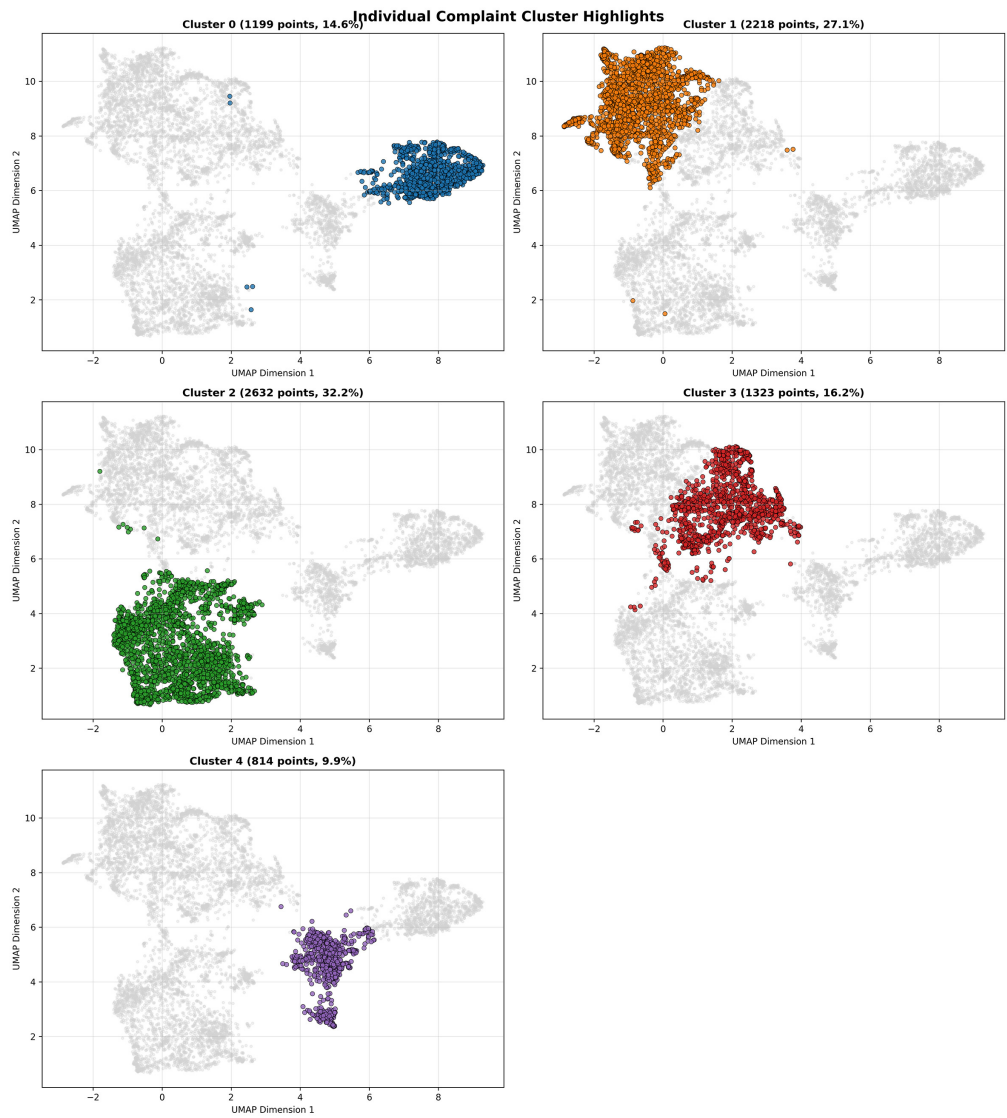
- “21년 8월이 넘으면 2년째입니다. 풀타임에서 반으로 줄고 계약서를 새로 안 씀. 새로 써야 하는데 이때 대표가 거부하거나 이전에 구두로 한 약속을 모른척하면 어떻게 하나요? 알바입니다.”
- “기업 법정 의무 교육 온라인 시 업무 특성상 근무 시간 내 교육이 불가하여 근무 외 시간 집 컴퓨터, 출퇴근 시 모바일, 기타 방법 등으로 교육할 경우 근로자 초과 근무 수당 지급 의무?”

클러스터 1 (교통과 도시환경 관련):

- “국토교통부에 문의드립니다. 이관하지 마시고 직접 답변 부탁드립니다... 평택 신도시 발코니 확장비 산정내역 공개 및 수정에 관한 민원 신청합니다.”
- “자동차배기가스에서 매연이 나오는 이유 문의. 매연이 너무 많이 나오는데 차에 문제가 있는 건가요? 왜 이렇게 많이 나오죠? 옆 차량 비교해도 눈에 확 차이가 나는데.”

클러스터 2 (부동산 및 건축 법규 관련 민원):

- “공동주택 신축공사 현장 사용승인 전 감리원료보고서(중간보고서) 작성주체 확인 요청의 건”



〈그림 5〉 민원 클러스터별 시각화

〈표 3〉 민원 클러스터 분포

| 클러스터 | 댓글 수(개) | 평균 문자 수 | 평균 단어 수 | 비율(%) |
|--------|---------|---------|---------|-------|
| 클러스터 0 | 1,199 | 121 | 29 | 14.6 |
| 클러스터 1 | 2,218 | 408 | 95 | 27.1 |
| 클러스터 2 | 2,632 | 394 | 88 | 32.2 |
| 클러스터 3 | 1,323 | 323 | 73 | 16.2 |
| 클러스터 4 | 814 | 173 | 40 | 9.9 |

- “막다른 도로에 대한 대지안의 공지 관련. 무더운 날씨에 노고가 많으십니다. 막다른 도로에 대한 대지 안의 공지 규정 관련입니다.”

클러스터 3 (행정 서비스 및 지원제도 관련 민원):

- “장애인주차구획을 확장형 주차구획으로 할 수 있는지? 안녕하십니까? 예를 들어 기계식 주차가 60대 있고 법적으로 장애인 주차를 2대 설치해야 되는 경우를 가정하고 여쭙보겠습니다.”
- “일용신고내역 정정 문제 문의. 안녕하세요. 일용직으로 일하는 중년 남성입니다. 다름이 아니오라 인력사무실에서 일하다가 구청에서 연락이 왔습니다.”

클러스터 4 (창업 및 경영 지원 관련 민원):

- “협동조합 설립 관련 자문. 울주군 소재 교육 서비스 업체입니다. 경기침체로 어려움을 겪고 있습니다. 최근 같이 일하시는 분들과 모여 협의하던 중 협동조합에 대한 의견이 많아서 요청을 드립니다.”
- “유망 기술개발 지원제도에 대하여. 중소기업을 운영 하고 있습니다. 기술을 개발을 위해 정부에서 지원해 주는 제도에 대하여 알고 싶습니다.”

유튜브 댓글 데이터에서는 4개 클러스터가 도출되었으며, 각 클러스터의 분포와 특성은 다음과 같다.

클러스터 0은 주로 부동산과 지역 관련 정보에 관심을 표하는 댓글로, 콘텐츠에 대한 평가나 감상을 정보 소비자 입장에서 표현하고 있었다. 평균 문자 수는 124.9자로 가장 긴 텍스트 길이를 보였으며, 이모지 사용이 28.98개였다. 반면 느낌표(0.26개)나 물음표(0.19개)의 사용은 상대적으로 적었으며, 감정을 주로 이

모지로 표현하는 특성을 보였다. 클러스터 1은 가장 큰 규모의 클러스터로 평균 73.3자의 문자 길이를 보였다. 느낌표의 사용이 0.58개로 적극적이고 감정적인 반응을 나타내는 반면, 물음표 사용은 0.04개로 가장 낮았다. 실제로 해당 클러스터는 개인적 경험과 감정을 상세히 공유하는 댓글 유형으로, 정서적인 반응과 감정을 비교적 길게 서술한 댓글이었다. 클러스터 2는 평균 문자 수가 70.2자로 클러스터 1과 유사한 길이를 보였다. 느낌표나 물음표 사용이 균형적으로 분포되어 다양한 감정 표현을 사용하는 특성을 보였다. 실제 데이터를 살펴보면 영상 제작자나 영상과 관련하여 친근한 표현으로 소통하는 댓글로, 대화지향적인 소통 양식이 나타났다. 클러스터 3은 가장 작은 규모의 클러스터로, 평균 26.1자로 가장 짧은 텍스트 길이를 보였다. 실제 간결한 감사를 나타내는 감정 표현을 사용하는 특성을 보였다. 각 클러스터의 대표적인 댓글 예시는 다음과 같다:

클러스터 0 (부동산 정보):

- “구미 원룸 시리즈!!!!!! 와 진짜 너무 사랑하는 구미 부산 원룸 시리즈 ㅋㅋㅋㅋ 보기 전에 댓글 달고 봅니다”
- “파이어족인 입장에선 구미영상은 그저 좋음 ㅎㅎ 다른 지역도 이리됐음 좋겠네”

클러스터 1 (개인 경험 공유):

- “꿈을 결정을 하지 못하고 방황하고 있었는데 원룸소년단 영상을 보면서 많은 위로와 용기를 얻어가요 아직도 뭐로 해먹고 살아야할지는 잘 모르겠지만 음악을 다시 시작하든, 노래를 하든, 미술을 하든 열심히 할 수 있을 것 같네요”



〈그림 6〉 유튜브 댓글 클러스터별 시각화

〈표 4〉 유튜브 댓글 데이터 클러스터 분포

| 클러스터 | 댓글 수 | 비율 | 평균 문자 수 | 평균 단어 수 | 느낌표 | 물음표 | 이모지 |
|--------|------|-------|---------|---------|------|------|-------|
| 클러스터 0 | 539개 | 27.0% | 124.9자 | 29.2자 | 0.26 | 0.19 | 28.98 |
| 클러스터 1 | 705개 | 35.4% | 73.3자 | 16.9자 | 0.58 | 0.04 | 16.60 |
| 클러스터 2 | 550개 | 27.6% | 70.2자 | 16.1자 | 0.45 | 0.13 | 15.98 |
| 클러스터 3 | 199개 | 10.0% | 26.1자 | 6.1자 | 0.41 | 0.05 | 6.09 |

- “맨마지막에 해주신 말이 되게 와닿네요... 혼자 자산다는건 나와 가까워지는 시간이다. 항상 타인에 대해서는 알기 위해 노력을 많이하는데... 정작 나 자신에 대해선 알려고 하지않는다는게 정말 공감되네요.”

클러스터 2 (친근한 소통):

- “안녕하세요! 바바방입니다!! 원뿔 구독자분들 관심 덕분에 이렇게 한 번 더 출연하게되

었네요!! 따뜻한 댓글 너무 감사드리고, 덕분에 촬영도 즐겁게 했습니다!!”

- “구미는 늘 새로워! 늘 짜릿해!”
- “요즘 내가 맨날 보는 OOOOOO는 재미있으면서도 마음도 즐겁고 시원해서 자주 보게돼용”

클러스터 3 (간결한 감사):

- “항상 응원하고 있어요! 좋은 영상 항상 올려 주셔서 감사합니다 덕분에 항상 웃네요ㅎㅎ”

- “수고많이했어요!!!오늘도 영상감사합니다 사랑해요”
- “좋은정보가 많네요”

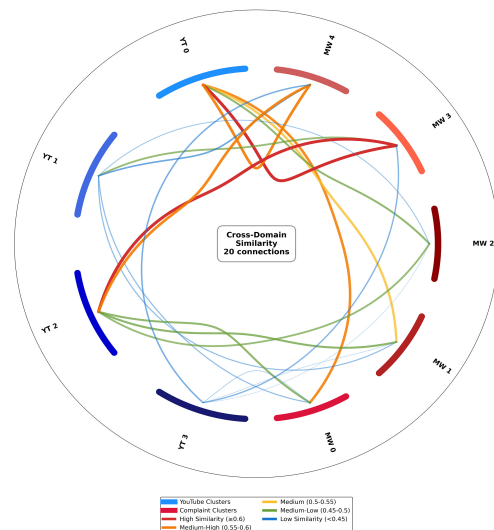
유튜브 댓글의 경우 클러스터 간 텍스트 길이의 차이, 이모티콘 사용 패턴, 감정 표현 방식에서 각기 다른 소통 특성이 나타났다. 모든 클러스터에서 ‘감사합니다’, ‘좋은’ 등의 긍정적인 표현이 공통적으로 등장했으나, 1인칭 대명사 사용 빈도와 감탄사, 맥락적 단어 조합에서는 클러스터 별 차이를 보였다.

4.2 잠재요구 세그먼트 분석

본 연구에서는 유튜브 댓글과 민원 데이터로 부터 도출된 클러스터들 간의 의미적 유사성을 측정하기 위하여 공통 관심 세그먼트를 도출하였다. 먼저 각 클러스터의 임베딩 벡터 중심점(centroid)을 계산하기 위해 클러스터에 속한 모든 문서의 임베딩 벡터를 평균하였다. 이후 유튜브 클러스터와 민원 클러스터 간 코사인 유사도를 산출하였다. 전체 클러스터 쌍의 유사도 분포를 분석한 결과와 선행연구를 바탕으로, 유사도 0.5 이상을 의미 있는 연관성, 0.6 이상을 높은 연관성으로 해석 기준을 설정하였다(Rekabsaz et al., 2017). 모든 클러스터 쌍을 유사도 순으로 정렬하여 상위 연관 쌍을 식별하고, 각 클러스터 쌍 별 평균 유사도와 상위 유사도를 지닌 클러스터 쌍을 분석하였다. 이를 통해 두 도메인 간 주제적 연결성이 높은 클러스터 쌍들을 공통 관심 세그먼트로 도출하였다.

〈그림 7〉은 유튜브 클러스터와 민원 클러스터 간 총 20개의 유사도 연결선을 표현해두었

으며, 연결선의 색상은 유사도 수준을, 연결선의 굵기는 유사도의 강도를 나타낸다. 유사도 범위가 0.247~0.630으로 분포하였고 평균 유사도는 0.442였다. 이 중 상위 5개 세그먼트를 주요 잠재 고객 세그먼트로 선정하였다.



〈그림 7〉 클러스터 간 유사도 시각화

도출된 상위 5개 세그먼트 유사도는 0.556~0.630 범위에 분포하여 높은 의미적 유사성을 보였다(〈표 5〉 참조). 이들 세그먼트는 각각 표현방식은 다르지만 본질적으로 유사한 정보 요구와 소통 패턴을 가지고 있음을 시사한다. 따라서 민원 클러스터와 유튜브 클러스터를 쌍으로 하여 하나의 통합 타겟 세그먼트로하여 마케팅 전략 수립에 활용 할 수 있다. 도출된 통합 잠재요구 세그먼트는 민원과 유튜브 댓글이라는 서로 다른 소통 창구에서 나타나는 유사한 정보 요구를 반영하고 있다. 이에 개인의 언어 사용이나 표현, 정서, 사회적 맥락, 심리적 동기를 다층적으로 반영하여 해석하고자 하였다.

〈표 5〉 유사도 기반의 통합 잠재 요구 세그먼트

| 민원 클러스터 | 유튜브 클러스터 | 유사도 | 통합 정보 요구 세그먼트 |
|---------|----------|-------|---------------|
| 3 | 0 | 0.629 | 정서 공감형 |
| 3 | 2 | 0.626 | 정보탐색형 |
| 4 | 0 | 0.579 | 정보 요구/소통형 |
| 4 | 2 | 0.576 | 제도 활용형 |
| 0 | 0 | 0.556 | 현실 참여형 |

4.3 논의

첫 번째 세그먼트는 유사도 0.629로 ‘정서 공감형’으로 명명하였다. 이 집단은 개인적인 경험과 지역 사회 공간에 대한 감정적 반응을 중심으로, 활발한 자기 표현과 타인과의 정서적 교류를 시도한다. 실제로 유튜브 데이터에서는 “꿈을 결정을 하지 못하고 방황하고 있었는데 00000 영상을 보면서 많은 위로와 용기를 얻어 가네요... 삶의 용기를 주셔서 너무너무 감사드립니다”, “서울은 진짜.. 헬이다..” 같은 감정적 토로가 나타났으며, 민원 데이터에서도 “저는 혼자 살고 있고요, 결혼도 못 했습니다... 해결할 방법이 없을까요?? ㅊㅊ”, “개인신용정보 유출이 걱정되어... 방법을 몰라 포기해 버렸다”처럼 개인적 상황을 상세히 드러내며 공감을 구하는 양상을 보인다. 이들은 활발한 자기 표현과 타인과의 정서적 교류를 시도하며, ‘나-지역’ 간의 정체성 형성과 공동체 의식이 내재되어 있다. 특히 “울집보다 썬데 저렇게 좋다니..34년후에 퇴직하면 지방으로 가야하나 봅니다”, “서울에 진짜 말도 안되는 집 너무 많음.. 애초에 작은 평수를 다 쪼개버리니 화장실 작고 세탁기 돌릴려면 신발장에서 신발 신고 돌리고”와 같이 지역에 대한 감정적 평가와 개인적 상황을 연결시키는 표현이 두드러진다.

이는 단순 정보 교환을 넘어서 정서적 연결망으로 확장하려는 동기가 뚜렷함을 시사한다. 즉 개인 경험과 지역사회에 대한 정서적 연결망을 중시하며, 사회적 소속감과 자기 정체성 강화의 욕구가 드러난다. 이는 사회적 정체성 이론에서 제시하는 집단 내 소속감 강화 행위와 부합하며 (Tajfel & Turner, 1986), Escalas(2004)의 연구에서 언급된 정서적 공감과 내러티브가 관계 형성에 미치는 영향과도 관련된다.

두 번째 세그먼트는 유사도 0.626로, ‘정보탐색형’이라 명명하였다. 해당 집단은 정책이나 행정 관련 정보를 체계적이고 목적 지향적으로 모으는 특성을 보인다. 유튜브에서는 “00000 Q&A 영상 만들어주세요!! 제가 궁금한건요, 1. 00000은 뭘이 세계관에서 잘생겼나요? 2.00000으로 번 돈 멤버들은 어떻게 사용했나요?”처럼 구체적이고 체계적인 질문을 제시하며, “00000 많은 응원 부탁드립니다”와 같이 적극적인 요청을 한다. 민원에서는 “장애인주차구획을 확장형 주차구획으로 할 수 있는지?... 확장형 규격에 맞게 설치해도 되는지 궁금합니다. 바쁘시겠지만 답변 부탁드립니다”, “기존 질의 회신을 볼 수 있는 방법은 어떠한 것이 있는지 궁금합니다” 등 명확한 정보와 해결책을 요구하는 양상을 보인다.

즉 정보에 대한 체계적 탐색과 합리적 의사

결정을 중시하는데, 이는 정보행동연구에서 탐색과 정보 활용 전략의 합리성에 대한 선행연구들(Kuhlthau, 1991; Wilson, 1999)과 그 맥을 같이 한다. 해당 연구들에 따르면 이러한 정보 탐색 활동은 불확실성을 줄이고 상황을 통제하려는 인지적 욕구에서 비롯된다. “구미 원룸 시리즈!!!!!! 와 진짜 너무 사랑하는 구미 부산 원룸 시리즈... 보기 전에 댓글 달고 봅시다”와 같이 정보 수집에 적극적으로 참여하는 모습을 보인다. Brashers(2001)의 불확실성 관련 이론에서 보면, 불확실성을 감소시키는 과정에서 정보의 신뢰성 및 정확성에 대한 요구가 높다. 이 세그먼트는 정보 수집과 검증에 능동적으로 참여하며, 정책 변화나 제도적 안내에 민감하게 반응한다. 이는 실용성이나 효율성 중심의 정보 이용 전략과 연결될 수 있다(Simon, 1979).

세 번째 세그먼트는 ‘정보 요구 및 소통형(유사도 0.626)’으로 명명하였다. 이 집단은 복잡한 내용을 이해하기 쉽게 간결하게 표현하는 것을 선호하며, 이에 긍정적 피드백을 제공하는 소통 스타일을 보인다. 유튜브에서는 “몰아서 보니 더 재밌네요오늘도 재밌는 영상 감사다”, “오늘도 잘보고 가요! 절대 접지마세여ㅜㅜ그럼 팟팅!!”처럼 간결하고 친근한 반응을 보인다. 민원에서는 “바쁘시겠지만 답변 부탁드립니다”, “빠른 답변 부탁드립니다” 등 명확하고 직접적인 요청 형태를 보인다. 정보의 가독성과 공감 가능성에 가치를 두는 집단으로, 언어적인 면에서는 평이하고 반복적이거나 간결한 구문이 많고, 대화형의 소통 형태를 보인다. “파이어족인 입장에선 구미영상은 그저 좋음 ㅎ 다른 지역도 이리됐음 좋겠네”와 같이 개인적 감상을

간결하게 표현하면서도 소통 의지를 드러낸다.

이는 미디어 심리학에서 메시지 가독성과 수용성의 중요성을 강조하는 여러 연구들과 관련이 있다(Hill & Levenhagen, 1995; Petty & Cacioppo, 1986). 이 연구들에서는 간결한 메시지가 정보 전달의 효율성을 높이고 청중의 참여도를 증진시킨다고 하였다. 또한 Luhmann(2000)은 간결한 커뮤니케이션이 사회적 신뢰 형성에 기여한다고 언급하였다. 간결한 커뮤니케이션 형식은 사회적 상호작용에서 신뢰구축과 협력적 관계 형성에 필요한 요소이다. 이들은 정보 요구가 높으면서도 정보를 쉽고 명확하게 이해하려는 ‘정보 수용자’이자, ‘소통 참여자’로 볼 수 있다.

네 번째 세그먼트는 유사도 0.576을 기록했으며, ‘제도 활용형’ 집단이다. 부동산이나 건축법규 등 제도적 또는 법률적 정보에 관심을 가지고 있다.

유튜브에서는 “파이어족인 입장에선 구미영상은 그저 좋음 ㅎ 다른 지역도 이리됐음 좋겠네”처럼 투자나 재정 관점에서 정보를 분석하고, “집공략 영상의 매력은 피디의 탄지를 진우형이 어떻게든 변명하는걸 보는게 재밌다는 것”과 같이 체계적 관찰 태도를 보인다. 민원에서는 “임대공장 50평의 공장등록은 어떻게 해야 하는지요?”, “수입 시 세관통관 시에 특별히 저축되는 것이 있는지요? 또는 FTA C/O를 발급받아 와야 하는지요?”, “폐업을 하게 되면 기존에 받은 사업자대출을 전액 상환해야 하나요?” 등 구체적인 절차와 규정에 대한 탐색적 질문을 한다.

Touraine(2000)은 사회학적 관점에서 제도를 설명하였는데, 제도란 개인과 집단의 권력 및 정체성을 규정하는 틀이며, 이 집단은 그 틀을 적

극적으로 활용하여 자신의 사회적 위치를 확보하려는 경향을 보인다. 이 집단은 전문성을 강조하는 표현과 합리성을 추구하는 표현들을 선호하는데, 이는 담론 분석 분야의 기술적 언어 사용으로 해석할 수 있다(Berger & Luckmann, 2016). 본 집단은 제도적인 맥락에서 자기 권리를 강화하려는 목적 지향적 행위자로, 정보의 정확성과 권위성을 중시한다.

마지막 세그먼트는 유사도 0.556으로, '현실 참여형'이다. 노동이나 근로 조건 관련 경험을 바탕으로 사회적 권리를 강화하고 이를 위해 실제적으로 실천적 참여를 추구하려는 정보 요구를 갖고 있다. 유튜브에서는 “이 분 나이만큼 자취한 저보다 더 나은거 같아요, 선정리 한거며 옷들들 말아서 세로 수납한것도 모자라... 역시 살림도 재능의 영역이 맞는거 같아요”, “진짜 인간 아무집 그자체”처럼 현실적 삶의 기술과 능력에 대한 평가를 하며, 민원에서는 “미사용 연차휴가 연차수당에 관련해 질문드립니다... 이번 년도에 못 쓴 거를 내년에 한 번에 다 쓸 수 있는지 궁금합니다”, “퇴사 후 회사가 1월 중 폐업하게 되고 퇴직금과 급여가 미지급된다면 어떻게 받아야 하나요? 최악의 경우 체당금으로 받을 수 있나요?” 등 구체적 권리 행사 방안을 묻는다.

이 집단은 불만을 표현하는 것에 그치지 않고, 자신의 경험을 토대로 사회 변화를 전환하고 싶어하는 적극성을 보인다. 언어적인 면에서는 강한 주체성이나, 변화를 요구하는 명령적 표현, 소속감을 강조하는 표현이 드러났다. Almond와 Verba(1963)와 Norris(2002)에서 언급하는 ‘행동적 시민성’으로 볼 수 있으며 강한 주체성과 집단 연대성을 언어로 표현하는 특성을 확인할 수 있었다(Polletta & Jasper, 2001).

5. 공공도서관 마케팅 적용 방안

앞서 도출된 다섯 개의 잠재 요구 세그먼트는 각각 차별화된 마케팅 접근 방향을 시사한다. 공공도서관이 제공하는 서비스가 단순 정보 전달을 넘어, 시민 삶의 방식과 가치를 반영하는 방식으로 재설계될 필요가 있음을 시사한다(Kotler & Lee, 2008). 특히 각 세그먼트는 시민의 심리적인 면과 사회적인 기대를 반영하고 있어, 향후 도서관 마케팅 전략 수립 시, 세분화(segmentation), 표적(targeting), 포지셔닝(positioning) 전략의 근거로 활용될 수 있다.

먼저 ‘정서 공감형’ 세그먼트는 지역 공간과 개인 경험 사이의 감정적 연결을 강조하는 특징을 가진다. 이 집단은 정보 소비를 넘어, 사회적 소속감과 자기정체성의 형성을 중요하게 인식하며, 이는 도서관 마케팅 전략에서 ‘감성 기반 가치 제안’의 가능성을 시사한다. 기존 연구에 따르면, 공공서비스 이용자는 정서적 만족과 사회적 연대감이 확보될 때, 기관에 대한 지속적인 관심과 긍정적인 태도를 형성하는 경향이 있다(Glückler & Ries, 2012; Berry, 2000). 특히 Berry(1999)는 감정적 유대 형성이 브랜드 신뢰와 관계 지속의 핵심이라 보고한 바 있으며, 이는 도서관과 이용자의 관계 형성에도 유의미하게 적용될 수 있다고 하였다. 또한 이 집단의 언어 표현을 보면, 자아 표현과 공동체 중심의 내러티브 표현이 나타났다. Escalas와 Bettman(2005)은 소비자의 자아정체성이 스토리텔링과 같은 자기서사 전략을 통해 강화된다고 밝힌 바 있으며, 이는 정보기관과 이용자 간의 관계에서 경험 공유 기반의 신뢰 형성 전략으로 확장이 가능하다고 하였다. 따라서 해당

세그먼트는 정보 중심 전략보다는 공감, 소속감, 경험의 사회적 가치를 중심으로 도서관의 정체성을 재구성할 수 있는 마케팅 대상 군집으로 볼 수 있다.

구체적인 서비스로는 지역 주민의 개인사와 지역사를 연결하는 구술사 프로젝트나 세대 간 경험 공유 플랫폼 제공, 개인의 생애주기별 맞춤형 정보 서비스, 개인의 감정적 상태와 관심사 변화를 추적하여 맞춤형 도서나 콘텐츠를 추천하는 큐레이션, 유사한 삶의 경험을 가진 시민들을 익명으로 연결하는 매칭 서비스, 개인의 다면적 정체성을 모두 지원하는 멀티 페르소나 프로그램 운영 등이 고려될 수 있다. 도서관은 정형적인 공지사항 대신 서사적인 언어를 활용한 소통을 통해 시민들의 일상적 고민과 감정에 공감하는 접점을 마련하고, 파편화된 시간을 의미 있게 연결하는 서비스와 '실패해도 괜찮은 공간, 완벽하지 않아도 환영받는 분위기' 조성을 통해 시민 개개인이 자신만의 성장을 지원받을 수 있는 정서적 안전망 역할을 수행하여야 한다.

‘정보 탐색형’ 세그먼트는 명확한 목적을 가진 정보 수집 활동과 정보의 신뢰성, 정확성에 관심을 갖는다. 이 집단은 합리적인 정보 행동을 기반으로 한 정보 접근 전략을 선호하며, 이는 Wilson(1997)의 정보추구 행동모델과 Kuhlthau(1991)의 정보 탐색 불확실성 이론과 부합한다. Kuhlthau(1991)의 연구에서는 정보탐색 초기 단계에서 이용자가 겪는 인지적 불안과 이를 해소하려는 전략적 정보행동의 중요성을 언급하였다. Savolainen(2015)은 디지털 환경에서의 정보행동도 불확실성 회피 전략이 중심에 있다고 보았으며, Schmidhuber et al.(2023)

의 연구에서는 정보 및 공공기관 이용자가 정보 정확성과 투명성을 신뢰의 핵심 요인으로 꼽았다. ‘정보 탐색형’ 세그먼트는 도서관이 정보의 신뢰성과 정합성, 검색 효율성을 핵심 가치로 제안할 수 있는 전략적 마케팅 대상이 될 수 있다.

가짜 뉴스와 정보 과부하가 심각한 현 시점에서 도서관은 팩트체크 서비스, 신뢰성이 검증된 정보원 데이터베이스 구축, 개인 맞춤형 정보 알림 서비스를 통해 기존의 신뢰할 수 있는 정보 큐레이터로서의 전문성을 강화할 수 있다. 특히 시간 효율성을 중시하는 현대인들을 위해서는 개인의 전문 분야에 대한 깊이 있는 자료 접근성을 높이고, 복잡한 정보를 신속하게 검증하고 활용할 수 있도록 지원하는 정보 리터러시 역량 개발 교육을 확대해야 한다.

‘정보 요구 및 소통형’ 세그먼트는 복잡한 정보를 단순화하고, 수용하기 쉬운 방식으로 재구성하는 커뮤니케이션에 가치를 두는 군집이다. 명료한 언어와 사회적 소통 가능성을 정보 소비의 핵심 기준으로 하며, 이는 Hill과 Levenhagen(1995) 연구의 메시지의 단순성, 명료성이 청중의 인지 부담을 줄여 정보의 수용성을 높인다는 결과와 부합한다.

도서관은 복잡한 행정정보나 전문정보를 시민이 이해하기 쉬운 형태로 번역하고 전달하는, 접근 가능한 정보 중개자 역할을 강화해야 한다. 구체적으로는 개인의 여러 사회적 역할을 지원하는 취향 기반의 커뮤니티 매칭 서비스나, ‘방해받지 않을 권리’를 보장하면서도 자연스러운 상호작용이 가능한 시민 간의 지식 공유 플랫폼, 그리고 세분화되는 개인의 관심사에 맞춘 소모임 지원 등을 통해 시민들이 도서관을 단순한

자료 이용공간으로 인식하기 보다는 ‘완전히 혼자, 완전히 함께도 아닌’ 적절한 거리감을 유지하는 사회적 연결의 거점으로 인식할 수 있도록 해야 한다.

‘제도 활용형’ 세그먼트는 법이나 정책 등 제도 기반의 정보에 대한 관심을 보였던 집단으로 해당 정보를 실질적으로 활용하고자 하는 목적 지향적 군집이다. 권위있는 출처, 전문적인 용어 활용, 정확한 문서화를 선호하며 이를 기반으로 신뢰를 형성한다. 전문정보 수요자로 분류할 수 있으며, 이들이 요구하는 정보환경은 일반 대중보다는 고차원적인 구조를 가진다(Vassilakaki et al., 2014). 최근 연구에서도 공공 서비스에서 제도 정보를 필요로 하는 정보 요구를 가진 이용자는 구체적이고 실행 가능한 정보 제공과 행정적인 신뢰 형성을 중요하게 여긴다고 밝혀진 바 있다(Bertot et al., 2012; Li & Feeney, 2014). 도서관은 해당 정보 요구 세그먼트에 대해 전문성을 기반으로 한 맞춤형 정보 큐레이션과 신뢰성 있는 자료 접근 서비스를 전략적으로 강화해야 한다. 구체적으로는 생활 밀착형 정책 정보 서비스나 개인 맞춤형 복지 혜택 안내, 창업 및 취업 관련 제도 정보 제공, 법률 상담 연계 서비스 등을 통해 시민의 실생활과 직결되는 제도적 정보에 대한 접근성을 획기적으로 개선할 수 있다. 특히 리빙랩 형태의 시민 참여형 정책 연구 지원을 통해 제도와 시민간의 간극을 줄이고, 권위 있는 출처와 정확한 문서화를 선호하는 이들의 신뢰를 구축하는 전문적 정보 파트너로서의 역할을 수행해야 한다.

마지막으로 ‘현실 참여형’ 세그먼트의 경우, 사회 구조에 대한 비판적 인식과 함께 실천적

인 참여 의지를 지닌 정보 요구 세그먼트이다. 정보를 수용하는 것을 넘어, 비판적 리터러시, 시민 참여적인 실천을 추구하는 군집으로, ‘참여적 시민’ 유형(Almond & Verba 1963)이나 ‘정보 주체로서의 시민’ 유형에 해당한다고 볼 수 있다(김지연, 2015). 최근에는 도서관이 사회참여를 위한 플랫폼으로 재조명되기도 하며, 정보 제공 기능을 넘어, 이용자의 실천 역량을 강화하는 공간으로 전환할 필요성이 제기되고 있다(Aabø, 2011). 이 세그먼트의 정보 요구를 해소하기 위해서는 도서관이 사회적 실천을 위한 정보 기반을 제공하는 지식 플랫폼으로서의 역할을 강화할 수 있다. 구체적으로 사회적 쟁점에 대한 객관적이고 다각적인 자료 큐레이션, 정책 분석 리포트와 시민 사회 연구 자료에 대한 접근 서비스 제공, 지역 현안 관련 데이터 및 통계 정보 서비스, 그리고 시민이 스스로 사회 문제를 분석하고 대안을 모색할 수 있도록 돕는 정보 분석 도구와 워크숍 운영 등을 통해 도서관을 단순한 정보 소비 공간이 아닌 사회적 실천을 위한 정보 거점으로 인식하게 할 수 있다.

세그먼트별 포지셔닝 전략과 차별화된 마케팅적 접근을 통해 공공도서관이 단일한 정체성에서 벗어나 다면적 가치 제안을 통해 다양한 시민 집단과의 관계를 구축할 수 있음을 시사한다. 본 연구에서 활용한 분석 방법론은 도서관 마케팅에서 기존의 사후적 대응 방식을 예측적 접근으로 전환할 수 있는 가능성을 제시하였다 이를 통해 각 세그먼트의 고유한 특성과 맥락을 반영한 맞춤형 서비스 경험을 설계할 수 있으며, 이는 단순히 개인화가 아니라 정보 요구의 심층적 동기와 사회적 맥락을 고려한 관계 구축 전략으로 발전하는 방향이다.

이러한 접근은 공공도서관의 사회적 역할을 단순한 정보 제공에서 시민의 삶의 질 향상을 위한 종합적 지원으로 확장시킬 수 있다. 각 세그먼트별로 차별화된 가치를 제공함으로써 도서관의 사회적 필요성과 고유성을 강화하고, 다양한 이해관계자로부터의 지지를 확보할 수 있다. 궁극적으로 도서관은 시민의 다층적 요구를 선제적으로 파악하고 이에 대응하는 능동적 파트너로 진화하여, 지역사회와의 연결을 강화하고 전통적인 독서 및 학습 지원 기능을 넘어서 시민의 일상적 정보 요구, 사회적 소통 욕구, 참여적 시민성 실현 등을 포괄적으로 지원하는 지역사회의 핵심 인프라로서의 역할을 확대할 수 있을 것이다.

6. 결 론

급변하는 사회 속에서 도서관이 시민의 삶에 지속적으로 기여하기 위해서는 서비스 기획의 전 과정에서 시민의 목소리를 반영하는 데이터 기반의 접근이 필요하다. 본 연구는 시민들의 요구를 민원 데이터와 유튜브 댓글이라는 비정형 데이터를 통해 분석함으로써, 도서관의 전략 수립에서 그간 간과되어 왔던 잠재 요구 발

굴이라는 과제를 외부 마케팅 관점에서 탐색하고자 하였다.

본 연구는 탐색적 연구로서 온라인상의 비정형 데이터 분석을 통해 공공도서관의 잠재수요 발굴의 가능성을 발견하고자 하였다. 민원 데이터와 유튜브 댓글 분석을 통해 정보 요구 기반의 세그먼트 도출하고, 각 세그먼트별 차별화된 마케팅적 접근이 필요함을 실증적으로 제시하였다. 이러한 접근은 도서관 서비스 기획자가 정보 요구의 특성에 기반한 서비스 대상 정의와 전략 구성을 가능하게 하였다.

그러나 본 연구는 민원 데이터와 유튜브 댓글 데이터의 양적 분석과 문헌 해석에 의존하여 공감적 접근법이 병행되지 못한 한계가 있으며, 또한 두 플랫폼으로 제한되어 다양한 정보 환경의 요구 패턴을 포괄하지 못할 수 있다는 한계가 있다. 이에 향후 연구에서는 다양한 디지털 플랫폼과 오프라인 환경을 포함한 확장된 데이터 수집을 통해 도서관 마케팅 세그먼트이션의 정밀도를 높이고, 시민 요구의 동적 변화에 대응하는 전략 수립 방안을 모색할 필요가 있다. 본 연구가 제시하는 프레임워크가 도서관이 변화하는 사회 속에서 시민들의 다양한 정보 요구에 보다 민감하고 적극적으로 대응하는데 기여할 수 있을 것으로 기대한다.

참 고 문 헌

- 강미희, 홍현진 (2012). 어린이도서관 공간특성과 이용행태에 관한 이용자 인식조사. 정보관리학회지, 29(1), 253-277. <https://dx.doi.org/10.3743/KOSIM.2012.29.1.253>
- 김미설, 장우권 (2015). 도서관 마케팅 프로모션의 연구 동향에 관한 연구. 정보관리학회지, 32(1),

- 171-204. <https://dx.doi.org/10.3743/KOSIM.2015.32.1.171>
- 김승재, 조규락 (2025). 감성분석을 활용한 유튜브 콘텐츠 학습자의 댓글 내용 연구. 컴퓨터교육학회 논문지, 28(2), 37-47. <https://doi.org/10.32431/kace.2025.28.2.004>
- 김지연 (2015). 빅데이터의 주제: 디지털 시민의 구성. 경제와 사회, 105, 94-126. <https://dx.doi.org/10.18207/criso.2015..105.94>
- 김형중, 김진화 (2017). 소셜 빅데이터 분석을 통한 소비자 가치 인식 연구: 신규 스마트폰을 중심으로. 한국전자거래학회지, 22(1), 123-146. <https://doi.org/10.7838/jsebs.2017.22.1.123>
- 박태연, 오효정 (2021). 뉴노멀 시대의 도서관 이용자 요구 분석: 소셜 미디어를 중심으로. 한국도서관·정보학회지, 52(2), 303-330. <http://doi.org/10.16981/kliss.52.2.202106.303>
- 신선아, 이명규 (2020). 영유아 보호자의 도서관 이용행태 및 재이용과 추천의사 영향 요인 연구: G광역시 공공도서관을 중심으로. 한국비블리아학회지, 31(4), 95-119. <https://dx.doi.org/10.14699/kbiblia.2020.31.4.095>
- 신호철, 허철무 (2022). 빅데이터 기반 머신러닝 기법을 활용한 소비자유형 분석 및 군집화에 관한 연구: 고객 리뷰평점 중심으로. 한국콘텐츠학회논문지, 23(2), 217-230. <https://doi.org/10.5392/JKCA.2023.23.02.217>
- 심지영 (2021). 도서대출 데이터를 이용한 공공도서관 이용자 정보요구 분석: 서울시 공공도서관을 중심으로. 정보관리학회지, 38(2), 173-199. <https://doi.org/10.3743/KOSIM.2021.38.2.173>
- 안성우 (2025). Mnet <스테이지 파이터> 유튜브 댓글 분석을 통한 대중 인식 연구. 한국무용과학회지, 42(2), 11-24. <https://doi.org/10.21539/Ksds.2025.42.2.11>
- 이고은, 김기영 (2014). 공공도서관의 이용자 민원 요인 분석 연구. 한국문헌정보학회지, 48(3), 281-301. <https://doi.org/10.4275/KSLIS.2014.48.3.281>
- 이성신 (2014). 도서관 마케팅 모형. 한국도서관·정보학회지, 45(3), 249-270. <https://doi.org/10.16981/kliss.45.3.201409.249>
- 이용재 (2012). 공공도서관 마케팅 전략: 뉴욕 공공도서관 사례를 중심으로. 한국도서관·정보학회지, 43(4), 113-137. <https://doi.org/10.16981/kliss.43.4.201212.113>
- 이정미 (2013). 빅데이터의 이해와 도서관 정보서비스에의 활용. 한국비블리아학회지, 24(4), 53-73. <http://dx.doi.org/10.14699/kbiblia.2013.24.4.053>
- 전채남, 서일원 (2013). 빅데이터 분석의 기술마케팅 활용에 관한 연구: 잠재 수요기업 발굴을 중심으로. 마케팅논집, 21(2), 181-203.
- 주가혜, 이현성, 김주연 (2023). 소셜 빅데이터 분석을 통한 장소마케팅 전략 모델에 관한 연구. 한국공간디자인학회논문집, 18(1), 313-325. <https://doi.org/10.35216/kisd.2023.18.1.313>
- 최민음, 오윤경, 정해린, 정다희 (2024). 버추얼 인플루언서 유튜브 콘텐츠에 대한 이용자 감정 반응 연구. 디지털콘텐츠학회논문지, 25(9), 2563-2574. <https://doi.org/10.9728/dcs.2024.25.9.2563>

- 최상희 (2021). 이용자 요구 기반 장서개발을 위한 대학도서관 희망도서 주제 분석. 한국문헌정보학회지, 58(1), 395-415. <http://dx.doi.org/10.4275/KSLIS.2024.58.1.395>
- 최지원, 박승진 (2024). 텍스트마이닝을 활용한 대전시 공공도서관 이용자의 인식과 경험 연구: SNS와 온라인 뉴스 기사를 중심으로. 한국문헌정보학회지, 58(2), 363-384.
<http://doi.org/10.4275/KSLIS.2024.58.2.363>
- 표순희, 김윤희, 김혜선, 김완중 (2015). 도서관 빅데이터 서비스 모형 개발에 관한 연구: 공공도서관을 중심으로. 정보관리학회지, 32(2), 63-86. <https://doi.org/10.3743/KOSIM.2015.32.2.063>
- 한계문 (2003). 도서관의 인터넷 마케팅을 위한 시장세분화 전략. 한국도서관·정보학회지, 34(1), 111-129.
- Aabø, S. (2011). The value of public libraries: a socio-economic analysis. In Belotti, M. (Ed.), *Verso Un'Economia Della Biblioteca: Finanziamenti, Programmazione e Valorizzazione in Tempo di Crisi*. Milano: Editrice Bibliografica, 169-176.
- Almond, G. A. & Verba, S. (1963). *The Civic Culture: Political Attitudes and Democracy in Five Nations*. New Jersey: Princeton University Press.
<http://www.jstor.org/stable/j.ctt183pnr2>
- Berger, P. & Luckmann, T. (2016). The social construction of reality. In *Social Theory Re-Wired*. Oxfordshire: Routledge, 110-122.
- Berry, L. L. (1999). *Discovering the Soul of Service: The Nine Drivers of Sustainable Business Success*. New York: Simon and Schuster.
- Berry, L. L. (2000). Cultivating service brand equity. *Journal of the Academy of Marketing Science*, 28, 128-137. <https://doi.org/10.1177/0092070300281012>
- Bertot, J. C., Jaeger, P. T., & Hansen, D. (2012). The impact of polices on government social media usage: issues, challenges, and recommendations. *Government Information Quarterly*, 29(1), 30-40. <https://doi.org/10.1016/j.giq.2011.04.004>
- Brashers, D. E. (2001). Communication and uncertainty management. *Journal of Communication*, 51(3), 477-497. <https://doi.org/10.1111/j.1460-2466.2001.tb02892.x>
- Bruce, H. L. (2010). *Identifying Hidden Consumer Needs: A Systematic Literature Review*. MRes Dissertation, Cranfield University, England.
- Chandratre, S. V. & Chandratre, M. S. (2015). Marketing of library and information services. *Journal of Commerce & Management Thought*, 6(1), 162-175.
<https://doi.10.5958/0976-478x.2015.000011.7>
- Das, S. (2012, March 1). *Discovering what people want before they do: building innovation platforms with latent consumer needs*. Point of View: Ipsos Marketing. Available:

- <https://www.ipsos.com/de-de/discovering-what-people-want-they-do>
- Escalas, J. E. & Bettman, J. R. (2005). Self-construal, reference groups, and brand meaning. *Journal of Consumer Research*, 32(3), 378-389. <https://doi.org/10.1086/497549>
- Escalas, J. E. (2004). Narrative processing: building consumer connections to brands. *Journal of Consumer Psychology*, 14(1-2), 168-180. https://doi.org/10.1207/s15327663jcp1401&2_19
- Eugenio-Martin, J. L. & Cazorla-Artiles, J. M. (2020). The shares method for revealing latent tourism demand. *Annals of Tourism Research*, 84, 102969. <https://doi.org/10.1016/j.annals.2020.102969>
- Germano, M. A. & Stretch-Stephens, S. M. (2012). Strategic value planning for libraries. *The Bottom Line: Managing Library*, 25(2), 71-88. <https://doi.org/10.1108/08880451211256405>
- Glückler, J. & Ries, M. (2012). Why being there is not enough: organized proximity in place-based philanthropy. *The Service Industries Journal*, 32(4), 515-529. <https://doi.org/10.1080/02642069.2011.596534>
- Goffin, K. & Lemke, F. (2004). Uncovering your customer's hidden needs. *European Business Forum*, 18, 45-47.
- Hill, R. C. & Levenhagen, M. (1995). Metaphors and mental models: sensemaking and sensegiving in innovative and entrepreneurial activities. *Journal of Management*, 21(6), 1057-1074. <https://doi.org/10.1177/014920639502100603>
- Islam, M. & Sheikh, S. A. (2024). Are social media-based marketing strategies the new mechanisms for attracting consumers? A quantitative method-based approach. *Journal of Theoretical and Applied Electronic Commerce Research*, 19, 3571-3583. <https://doi.org/10.3390/jtaer19040173>
- Kotler, P. & Armstrong, G. (2020). *Principles of Marketing* (18th edition). London: Pearson Education Limited.
- Kotler, P. & Lee, N. (2008). *Social Marketing: Influencing Behaviors for Good*. Los Angeles: Sage Publications.
- Kuhlthau, C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), 361-371. [https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5<361::AID-ASI6>3.0.CO;2-%23](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-%23)
- Kumber, R. D. (2004). The importance marketing and total quality management in libraries. *Electronic Journal of Academic and Special Librarianship*, 5(2-3), 1-1.
- Li, M. & Feeney, M. (2014). Adoption of electronic technologies in local US governments: distinguishing between e-services and communication technologies. *The American Review*

- of Public Administration, 44(1), 75-91. <https://doi.org/10.1177/0275074012460910>
- Luhmann, N. (2000). Familiarity, confidence, trust: problems and alternatives. In Diego, G. eds. Trust: Making and Breaking Cooperative Relations, 6, 94-107.
- McInnes, L., Healy, J., & Melville, J. (2018). Umap: uniform manifold approximation and projection for dimension reduction. Journal of Open Source Software, 3(29), 861. <https://doi.org/10.21105/joss.00861>
- Mu, J., Bao, Y., Sekhon, T., Qi, J., & Love, E. (2018). Outside-in marketing capability and firm performance. Industrial Marketing Management, 75, 37-54. <https://doi.org/10.1016/j.indmarman.2018.03.010>
- Musarra, G. & Morgan, N. A. (2020). Outside-in marketing: renaissance and future. Industrial Marketing Management, 89, 98-101. <https://doi.org/10.1016/j.indmarman.2020.01.001>
- Norris, P. (2002). Democratic Phoenix: Reinventing Political Activism. Cambridge: Cambridge University Press.
- Odung, R. & Agungi, B. (2023). Marketing strategies and user appreciation of the services provided by public libraries. Library Philosophy and Practice (e-journal), 7910, 1-16. <https://digitalcommons.unl.edu/libphilprac/7910>
- Park, S. & Shin, H. (2021). KR-SBERT: a pre-trained Korean-specific sentence-BERT. GitHub Repository. Available: <https://github.com/snunlp/KR-SBERT>
- Petty, R. & Cacioppo, J. (1986). The elaboration likelihood model of persuasion. Advances in Experimental Social Psychology, 19, 123-205. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Polletta, F. & Jasper, J. M. (2001). Collective identity and social movements. Annual Review of Sociology, 27(1), 283-305. <https://doi.org/10.1146/annurev.soc.27.1.283>
- Reimers, N. & Gurevych, I. (2019). Sentence-bert: sentence embeddings using siamese bert-networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, 3982-3992.
- Rekabsaz, N., Lupu, M., & Hanbury, A. (2017). Exploration of a threshold for similarity based on uncertainty in word embedding. Paper presented at European Conference on Information Retrieval, 10193, 396-409. https://doi.org/10.1007/978-3-319-56608-5_31
- Rosário, A. T. & Dias, J. C. (2023). How has data-driven marketing evolved: challenges and opportunities with emerging technologies. International Journal of Information Management Data Insights, 3, 1-14. <https://doi.org/10.1016/j.jiime.2023.100203>
- Rust, R. T. (2020). Outside-in marketing: why, when and how?. Industrial Marketing Management,

- 89, 102-104. <https://doi.org/10.1016/j.indmarman.2019.12.003>
- Satish, K. V. R. & Kavya, N. P. (2016). Trend analysis of e-commerce data using hadoop ecosystem. *International Journal of Computer Applications*, 147(6), 1-5. <https://doi.org/10.5120/ijca2016911109>
- Savolainen, R. (2015). Cognitive barriers to information seeking: a conceptual analysis. *Journal of Information Science*, 41(5), 613-623. <https://doi.org/10.1177/0165551515558785>
- Schmidhuber, L., Willems, J., & Krabina, B. (2023). Trust in public performance information: the effect of data accessibility and data source. *Public Administration Review*, 83(2), 279-295. <https://doi.org/10.1111/puar.13603>
- Simon, H. A. (1979). Information processing models of cognition. *Annual Review of Psychology*, 30(1), 363-396. <https://doi.org/10.1146/annurev.ps.30.020179.002051>
- Styf, Y. (2020). The Relation Between Information Management Strategies and an Organisation's Marketing Strategy - A Comparative Study of Three Logistics Companies. Master thesis, The Swedish School of Library and Information Science, Borås.
- Tajfel, H. & Turner, J. C. (1986). The social identity theory of intergroup behavior. In Worchel, S. & Austin, W.G., eds. *Psychology of Intergroup Relation*, Chicago: Hall Publishers, 7-24.
- Touraine, A. (2000). A method for studying social actors. *Journal of World-Systems Research*, 6(3), 900-918.
- Varadarajan, R. (2020). Customer information resources advantage, marketing strategy and business performance: a market resources based view. *Industrial Marketing Management*, 89, 89-97. <https://doi.org/10.1016/j.indmarman.2020.03.003>
- Vassilakaki, E., Garoufallou, E., Johnson, F. & Hartley, R. J. (2014). Users' information search behavior in a professional search environment: a methodological approach. In Paltoglou, G., Loizides, F. & Hansen, P. eds. *Professional Search in the Modern World*. Lecture Notes in Computer Science, Cham: Springer, 23-44. https://doi.org/10.1007/978-3-319-12511-4_3
- Wilson, T. D. (1997). Information behaviour: an interdisciplinary perspective. *Information Processing & Management*, 33(4), 551-572. [https://doi.org/10.1016/S0306-4573\(97\)00028-9](https://doi.org/10.1016/S0306-4573(97)00028-9)
- Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249-270. <https://doi.org/10.1108/EUM0000000007145>
- Zou, H., Chen, H. M., & Dey, S. (2020). Engaging users through social media in public libraries. *ArXiv*. <https://doi.org/10.48550/arXiv.2003.04204>

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- Ahn, Sungwoo (2025). A study on public perception through YouTube comment analysis of MNET stage fighter. *Official Journal of Korean Society of Dance Science*, 42(2), 11-24. <https://doi.org/10.21539/Ksds.2025.42.2.11>
- Choi, Jiwon & Kwak, Seung-Jin (2024). A study on the perception and experience of Daejeon public library users using text mining: focusing on SNS and online news articles. *Journal of the Korean Society for Library and Information Science*, 58(2), 363-384. <http://doi.org/10.4275/KSLIS.2024.58.2.363>
- Choi, Mideum, Oh, Yun Kyung, Cheong, Hae Rin, & Chung, Dahee (2024). Emotional response to virtual influencer content on YouTube. *Journal of Digital Contents Society*, 25(9), 2563-2574. <https://doi.org/10.9728/dcs.2024.25.9.2563>
- Choi, Sanghee (2021). A topic analysis of requested books by user types at a university library for patron-derived acquisition. *Journal of the Korean Society for Library and Information Science*, 58(1), 395-415. <http://dx.doi.org/10.4275/KSLIS.2024.58.1.395>
- Han, Kay-Moon (2003). Market segmentation strategy for internet marketing in libraries. *Journal of Korean Library and Information Science Society*, 34(1), 111-129.
- Jun, ChaeNum & Seo, Il Won (2013). Analyzing the bigdata for practical using into technology marketing: focusing on the potential buyer extraction. *Journal of Marketing Studies*, 21(2), 181-203.
- Kang, Mi-Hee & Hong, Hyun-Jin (2012). Survey on the user's recognition about spatial characteristics and user behaviors in the children's library. *Journal of the Korean Society for Information Management*, 29(1), 253-277. <https://dx.doi.org/10.3743/KOSIM.2012.29.1.253>
- Kim, Hyong-Jung & Kim, Jin-Hwa (2017). A study on consumer value perception through social big data analysis: focus on smartphone brands. *The Journal of Society for e-Business Studies*, 22(1), 123-146. <https://doi.org/10.7838/jsebs.2017.22.1.123>
- Kim, Ji Yeon (2015). Big data & its subjects - digital citizens are constructed. *Economy and Society*, 105, 94-126. <https://dx.doi.org/10.18207/criso.2015..105.94>
- Kim, Mi-Seol & Chang, Woo-Kwon (2015). A study on the research trend of library marketing promotion. *Journal of the Korean Society for Information Management*, 32(1), 171-204. <https://dx.doi.org/10.3743/KOSIM.2015.32.1.171>
- Kim, Seungjae & Cho, Kyoo-Lak (2025). A study on YouTube learner comments using sentiment analysis. *The Journal of Korean Association of Computer Education*, 28(2), 37-47.

- <https://doi.org/10.32431/kace.2025.28.2.004>
- Lee, Goeun & Kim, Giyeong (2014). A study on the factors of civil petitions & complaints within public libraries. *Journal of the Korean Society for Library and Information Science*, 48(3), 281-301. <https://doi.org/10.4275/KSLIS.2014.48.3.281>
- Lee, Jeong-Mee (2013). Understanding big data and utilizing its analysis into library and information services. *Journal of the Korean BIBLIA Society for Library and Information Science*, 24(4), 53-73. <http://dx.doi.org/10.14699/kbiblia.2013.24.4.053>
- Lee, Seongsin (2014). Library marketing model. *Journal of Korean Library and Information Science Society*, 45(3), 249-270. <https://doi.org/10.16981/kliss.45.3.201409.249>
- Lee, Yong Jae (2012). A study on marketing strategies of public libraries: focused on New York public libraries. *Journal of Korean Library and Information Science Society*, 43(4), 113-137. <https://doi.org/10.16981/kliss.43.4.201212.113>
- Park, Tae-Yeon & Oh, Hyo-Jung (2021). Analysis of library user needs in the new normal era: focusing on social media. *Journal of Korean Library and Information Science Society*, 52(2), 303-330. <http://doi.org/10.16981/kliss.52.2.202106.303>
- Pyo, Soonhee, Kim, Yun Hyung, Kim, Hyesun, & Kim, Wan Jong (2015). A study on the developing of big data services in public library. *Journal of the Korean Society for Information Management*, 32(2), 63-86. <https://doi.org/10.3743/KOSIM.2015.32.2.063>
- Shim, Jiyoung (2021). Identifying information needs of public library users based on circulation data: focusing on public libraries in Seoul. *Journal of the Korean Society for Information Management*, 38(2), 173-199. <https://doi.org/10.3743/KOSIM.2021.38.2.173>
- Shin, Hocheol & Heo, Chul Moo (2023). A study on the analysis and clustering of consumer types using big data-based machine learning techniques: focus on customer reviews. *Journal of the Korea Contents Association*, 23(2), 217-231. <https://doi.org/10.5392/JKCA.2023.23.02.217>
- Shin, Seon-A & Lee, Myounggyu (2020). A study on infant caregivers' library use behavior and factors affecting their revisit and intention of recommending library visit to other people: focusing on public library services in G-city. *Journal of the Korean BIBLIA Society for Library and Information Science*, 31(4), 95-119. <https://dx.doi.org/10.14699/kbiblia.2020.31.4.095>
- Zhu, Jiahui, Lee, Hyunsung, & Kim, Joo Yun (2023). A study on the place marketing strategy elements based on social big data. *Journal of Korea Institute of Spatial Design*, 18(1), 313-325. <https://doi.org/10.35216/kisd.2023.18.1.313>