

<http://dx.doi.org/10.17703/JCCT.2022.8.6.703>

JCCT 2022-11-86

이미지 검색: 정보과다 환경에서의 접근과 이용

Image Retrieval: Access and Use in Information Overload

박민수*

Minsoo Park*

요약 학술문헌의 표와 그림에는 중요하고 가치 있는 정보가 포함되어 있다. 표와 그림은 정제된 연구의 본질을 나타내며 이는 원시 데이터셋에 가장 가까운 것이라 할 수 있다. 그렇다면, 연구자들은 검색시스템을 통하여 이러한 이미지 데이터에 쉽게 접근하여 활용할 수 있는가? 본 연구에서는 이용자연구 문헌조사와 국내외 사례조사 분석을 통하여, 이미지 데이터에 대한 이용자 인식 및 니즈를 파악하고 이미지 검색시스템에 대한 잠재적인 기대효과 및 활용 방안을 모색해보고자 한다. 대다수의 연구자들은 표 및 그림 색인 기능과 기존 검색 기능을 결합한 시스템을 선호하는 것으로 나타났다. 특정 개체 유형(그림 및 표)으로 검색을 제한할 수 있는 고급 검색 기능의 제공을 매우 중요하게 평가했다. 이와 관련하여, 그림과 표에 대한 검색 제한의 구현에 가장 높은 만족도를 주었다. 전반적으로, 연구자들은 표와 그림을 색인화하는 시스템의 많은 잠재적 용도를 발견할 수 있었다. 교육, 발표, 연구 및 학습을 위한 정보와 특수한 유형의 정보를 찾는 데 도움이 될 수 있는 것으로 나타났다. 이러한 시스템의 유용성은 기능이 기존 시스템에 통합되고 폴텍스트에 원활하게 연결되며 완전한 캡션이 있는 고품질 이미지를 포함하는 경우 가장 높게 나타났다. 이용자 중심 이미지 검색시스템에 대한 기대효과와 활용방안 또한 논의되었다.

주요어 : 이미지 검색, 이용자연구, 사례조사, 정보시스템, 과학기술

Abstract Tables and figures in academic literature contain important and valuable information. Tables and figures represent the essence of the refined study, which is the closest to the raw dataset. If so, can researchers easily access and utilize these image data through the search system? In this study, we try to identify user perceptions and needs for image data through user and case studies. Through this study we also explore expected effects and utilizations of image search systems. It was found that the majority of researchers prefer a system that combines table and figure indexing functions with traditional search functions. They valued the provision of an advanced search function that would allow them to limit their searches to specific object types (pictures and tables). Overall, researchers discovered many potential uses of the system for indexing tables and figures. It has been shown to be helpful in finding special types of information for teaching, presentation, research and learning. It should be also noticed that the usefulness of these systems is highest when features are integrated into existing systems, seamlessly link to fulltexts, and include high-quality images with full captions. Expected effects and utilizations for user-centered image search systems are also discussed.

Key words : Image Retrieval, User Studies, Case Studies, Information System, Science and Technology

*정희원, 강남대학교 데이터사이언스전공, 부교수 (단독저자)
접수일: 2022년 9월 26일, 수정완료일: 2022년 10월 25일
게재확정일: 2022년 11월 1일

Received: September 26, 2022 / Revised: October 25, 2022
Accepted: November 1, 2022

*Corresponding Author: mspark7@gmail.com
Dept. of Data Science, Kangnam Univ, Korea

I. 서론

학술문헌의 표와 그림에는 중요하고 가치 있는 정보가 포함되어 있다. 표와 그림은 정제된 연구의 본질을 나타내며 이는 원시 데이터셋에 가장 가까운 것이라 할 수 있다. 그렇다면, 연구자들은 검색시스템을 통하여 이러한 이미지 데이터에 쉽게 접근하여 활용할 수 있는가?

연구자들은 표, 그림, 그래프보다는 사진과 지도를 검색하며, 주로 구글 이미지 검색기능을 활용한다. 이러한 검색엔진을 통한 일반적인 검색에 대한 만족도는 지속적으로 낮은 평가를 받고 있다 [1-3]. 왜냐하면 일반적으로 학술적 표, 그림, 그래프는 '숨겨진 데이터(hidden data)'로 검색에 제한이 있기 때문이다. 다시 말해, 검색이 불가능하거나 매우 어렵다는 것을 의미한다.

따라서 본 연구에서는 이용자연구 문헌조사와 국내 사례조사 분석을 통하여, 이미지 데이터에 대한 이용자 인식 및 니즈를 파악하고 이미지 검색시스템에 대한 잠재적인 기대효과 및 활용방안을 모색해보고자 한다.

II. 이미지 검색: 이용자 연구

이미지 이용자와 관련 있는 많은 연구들이 이미지 검색의 유용성을 테스트하고 이미지 검색이 연구 활동에 어떤 영향을 미치는지를 조사하였다 [4-11]. 다양한 기관들에서 과학기술분야 연구자들이 연구에 참여하였다. 대표적인 연구의 질문은 다음과 같다: (1) 표와 그림의 색인이 연구자들에게 어떻게 이용될 수 있는가?, (2) 표와 그림 시스템에 특히 유용한 검색과 인터페이스 기능들은 어떤 것들이 있을 수 있는가? 설문지, 관찰, 구조화된 다이어리 등 복수 연구 도구들을 사용하여 데이터를 수집하였다. 표와 그림 색인 프로토타입 데이터베이스가 활용되었다.

대다수의 연구자들은 교육과 연구 모두에서 고품질의 그래프, 그림 또는 표에 대한 접근의 중요성을 언급하였으며, 소수의 참여자들만이 웹 검색엔진을 통해 지도나 사진을 주로 검색하는 것으로 나타났다. 대부분의 연구자들은 이미지 데이터가 전체 아티클(fulltexts)의 문맥 내에서 가장 유용하다고 느꼈고 일부는 맥락 없이 이미지를 이용하는 것에 대해 주의를 기울였다. 대부분의 경우 완전한 캡션, 표 또는 그림의 고품질 복제 및

폴텍스트에 대한 링크가 중요한 것으로 나타났다. 동일한 정보를 찾는 데 있어서, 전문 데이터베이스에서 더 오랜 검색시간이 소요될 것이라고 언급되었다.

이들 연구에 따르면, 특정 유형의 정보 검색기능은 이용자의 검색 프로세스에 차이를 가져올 수 있다. 예를 들면, 시간을 절약하고, 더 효율적으로 작업하고, 프레젠테이션을 지원하고, 더 관련성 있는 결과를 찾을 수 있다 [12-14].

대다수의 연구자들은 표 및 그림 색인 기능과 기존 검색 기능을 결합한 시스템을 선호하는 것으로 나타났다. 특정 개체 유형(그림 및 표)으로 검색을 제한할 수 있는 고급 검색 기능의 제공을 매우 중요하게 평가했다. 이와 관련하여, 그림과 표에 대한 검색 제한의 구현에 가장 높은 만족도를 주었다.

전반적으로, 연구자들은 표와 그림을 색인화하는 시스템의 많은 잠재적 용도를 발견할 수 있었다. 교육, 발표, 연구 및 학습을 위한 정보와 특수한 유형의 정보를 찾는 데 도움이 될 수 있는 것으로 나타났다. 이러한 시스템의 유용성은 기능이 기존 시스템에 통합되고 폴텍스트에 원활하게 연결되며 완전한 캡션이 있는 고품질 이미지를 포함하는 경우 가장 높게 나타났다.

Elsevier의 SciVerse ScienceDirect Image Search를 대상으로 한 이미지 이용자 연구에 따르면, 정보 검색 속도는 연구자들의 워크플로우 향상과 밀접한 관계에 있다 [15-17]. 과학기술 분야 연구자들을 대상으로 포커스그룹인터뷰를 수행한 결과, 많은 연구자들이 이미지 검색하고자 할 때 웹 검색엔진에 의존하는 것으로 나타났다(예를 들면, 구글의 이미지 검색 기능). 이미지 데이터와 관련 논문들이 함께 검색이 가능할 때 검색 결과에 대한 적합성을 평가하는데 도움이 되는 것으로 나타났다.

구글 이미지 검색을 자주 사용하는 연구자들은, 정보의 완결성과 신뢰성을 전문 데이터베이스의 차별성으로 평가하고 있다. 예를 들면, 전문 데이터베이스에서 제공하는 이미지 자료의 출처와 참조를 알 수 있는 반면, 웹 검색엔진 상에서 검색한 정보는 누군가 다른 곳에서 차용했을 수 있고 원본 출처에 적절한 크레딧이 제공되지 않을 수 있다.

이 외에도, 시각적 학습 기술을 선호하는 사람들에게 이미지 검색이 큰 이점이 있을 수 있다 [4]. 이미지 검색은 한 장의 사진이 천 마디 말의 가치를 전달할 수

있기 때문이다.

III. 사례 조사

KORDI (한국해양연구원) 사진 DB 서비스 (그림 1)는 정부기관 최초로 해양과학 사진 데이터베이스를 구축하여 대국민 서비스를 운영하고 있다. 다양한 바닷속 사진과 연구사업 수행 중 생산한 해양과학 사진자료로 구성한다 [18].



그림 1. KORDI 사진 데이터베이스
 Figure 1. KORDI Image Database

KAERI (한국원자력연구원) 이미지 DB 서비스 (그림 2)는 원자력 분야 논문, 특히, 강의자료, 백과사전 등 다양한 콘텐츠 내 이미지를 서비스하고 있다. 검색 항목으로는 서명, 키워드, 해설, 자료출처로 검색이 가능하다 [19].

KOSFIC (해양수산연구정보센터) 학술지 이미지 DB 서비스 (그림 3)는 국내 학술지 논문의 그림과 표 이미지를 제공한다. 원문 연계 서비스 또한 제공하고 있다. 그림 캡션, 조사해역, 키워드, 저자로 검색이 가능하다 [20].



그림 2. KAERI 이미지 데이터베이스
 Figure 2. KAERI Image Database

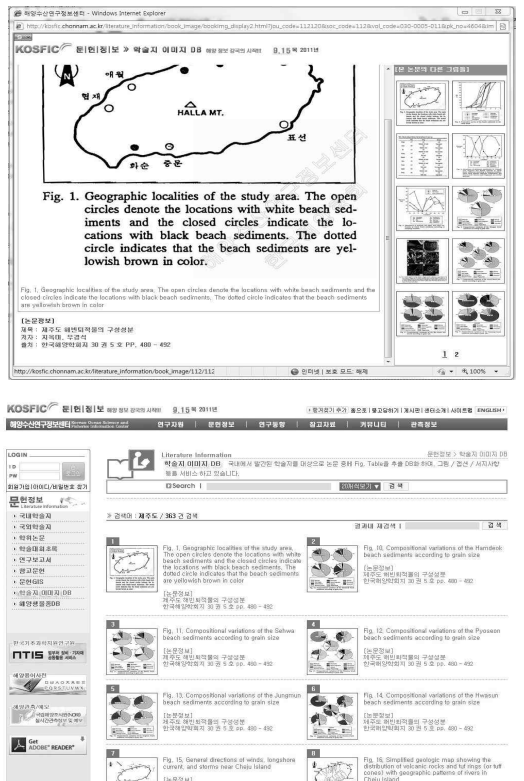


그림 3. KOSFIC 이미지 데이터베이스
 Figure 3. KOSFIC Image Database

PubMed의 이미지 데이터베이스 (그림 4)는 의학, 생명과학 분야 학술지의 이미지와 도표를 제공한다. PubMed Central을 통해 이미지 검색서비스가 제공되고 있다. 그림이나 표의 캡션으로 제한 검색이 가능하다 [21].

Biology Image Library (그림 5)는 BioMed Central 이 운영하는 생물학, 생의학 분야 온라인 이미지 컬렉션이다. 이미지 뿐 아니라 동영상, 삽화, 애니메이션 등 다양한 형태로 제공한다. BioMed Central에서 논문 내 이미지를 제공하고 있다 [22].

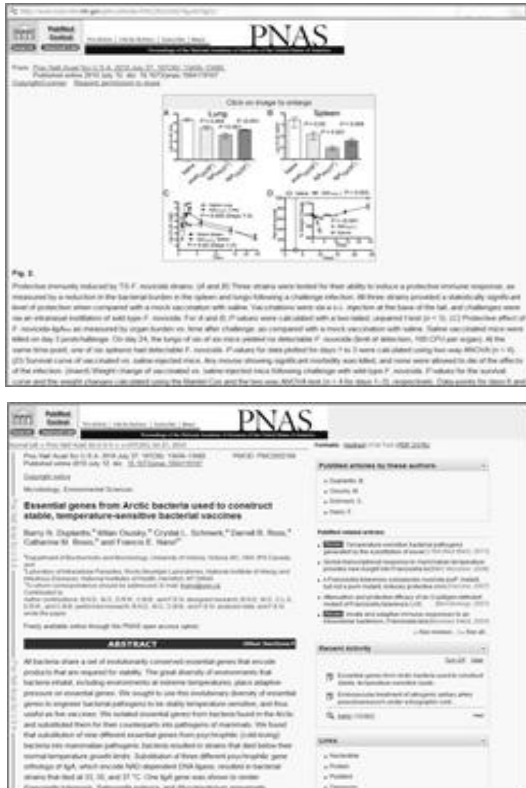


그림 4. PubMed 이미지 데이터베이스
Figure 4. PubMed Image Database



그림 5. Biology Image Library
Figure 5. Biology Image Library

IV. 결 론

본 연구에서는, 관련 문헌조사와 국, 내외의 사례조사를 통해, 이미지 검색에 대한 연구자들(잠재적 이용자를 포함한)의 인식과 니즈를 조사하였다. 주요 이슈 들로는 다음과 같다: (1) 정보 검색 및 조직 기술, (2) 대규모 데이터 처리를 가능하게 하는 인프라, (3) 인간-시스템 상호 작용의 문제, (4) 사회적 문제.

첫째, 사용자 질의어를 가장 관련성 높은 이미지와 매칭시키기 위해 보다 향상된 이미지 검색 방법론이 필요하다. 정보의 폭발은 검색 결과의 재현율을 상승시키고, 이로 인해 검색자는 필요로 하는 자료에 대한 판단에 소요되는 시간은 늘어날 수 있는 상황에 놓일 수 있다. 수많은 검색 결과 가운데 가장 관련성이 높은 이미지를 최소한의 노력으로 찾을 수 있어야 할 것이다.

둘째, 이러한 고급 방법론은 강력한 컴퓨터 시스템을 필요로 한다. 이미지는 텍스트보다 크기가 크고 조작이 훨씬 어렵기 때문에 이를 관리할 수 있는 인프라를 구축할 필요가 있다. 이미지 검색 시스템의 확장성, 안정성 및 속도가 주요 이슈들이다.

셋째, 사람들이 이미지를 사용하는 방식도 변화하고 있다. 이미지 매체는 더 이상 제한된 수의 전문가만을 위한 것이 아니다. 거의 모든 사람이 사용한다고 볼 수 있다. 기존의 이러한 변화에 상응하기 위해, 새로운 정보 행동을 연구하고 그 결과를 시스템 설계에 대한 과거 사용자 연구에서 축적된 지식과 통합할 필요가 있다.

넷째, 이미지 매체의 지속적인 대중화는 법적 갈등과 문화적 차이 등과 같은 다양한 사회적 문제를 염두에 둘 필요가 있다.

연구자 측면에서 바라본 기대효과는 다음과 같다: 우선, 방대한 정보 속 필요한 부분에 대한 보다 신속하고 정확한 검색이 가능하다. 이미지 출처(원문)로의 용이한 접근 및 상세 정보에 대한 검색이 가능하다. 이미지 배경에 대한 빠른 이해와 문헌 리뷰에 필요한 시간을 대폭 줄일 수 있다. 최신 데이터와 연구 결과에 대한 접근이 가능하다. 개인화서비스 또는 검색 최적화를 통해 개인의 정보 요구에 따른 검색 환경의 설정이 가능하다.

전문 정보를 제공하는 정보 전문가 측면에서 바라본 활용방안은 다음과 같다: 연구자들의 연구 생산성을 높이는데 도움을 줄 수 있다. 이는 문헌 리뷰에 소요되는 시간을 대폭 줄일 수 있기 때문이다. 최신 데이터의 지속적인 제공이 가능하다. 신뢰할 수 있는 출처로부터의 고품질 데이터를 제공할 수 있다. 이는 웹 검색엔진과의 차별화로 작용할 수 있다. 잠재적 이용자를 포함한 일반 이용자들은 보다 쉽고 빠르게 이미지를 검색할 수 있다. 이미지 서비스에 대한 축적된 이용자 모니터링 데이터를 통해 이미지 검색시스템 기능을 지속적으로 개선해 나갈 수 있다.

References

- [1] M. Park and T. Lee, "Understanding science and technology information users through transaction log analysis," *Library Hi Tech*, Vol. 31, No. 1, pp. 123-140, 2013.
- [2] M. Park and T. Lee, "A longitudinal study of information needs and search behaviors in science and technology: A query analysis," *Electronic Library*, Vol. 34, No. 1, pp. 83-98, 2016.
- [3] M. Park, "In-depth understanding of STEM information needs using FGI," *International Journal of Advanced Culture Technology*, Vol. 8, No. 3, pp.280-284, 2020.
- [4] M. Park, "Cognitive factors in adaptive information access," *International Journal of Advanced Culture Technology*, Vol. 6, No. 4, pp. 309-316, 2018.
- [5] M. Inoue, "Image retrieval: Research and use in the information exploration," *Progress in Informatics*, No. 6, pp.3-14, 2009.
- [6] R. Sandusky and C. Tenopir, "Finding and using journal article components: Impacts of disaggregation on teaching and research practice," *Journal of the American Society for Information Science and Technology*, Vol. 59, No. 6, pp.970-982, 2008.
- [7] P. Enser, "The evolution of visual information retrieval," *Journal of Information Science*, Vol. 34, No. 4, pp.531-546, 2008.
- [8] L. H. Armitage and P. G. Enser, "Analysis of user need in image archives," *Journal of Information Science*, Vol. 23, No. 4, pp.287-299, 1997.
- [9] C. Tenopir, P. Wang, Y. Zhang, B. Simmons and R. Pollard, "Academic users' interaction with ScienceDirect in search tasks: Affective and cognitive behaviors," *Information Processing & Management*, Vol. 44, No. 1, pp. 105-121, 2008.
- [10] I. Rowlands et al., "The Google generation: the information behavior of the research of the future," *Aslib Proceedings*, Vol. 60, No. 4, pp. 290-310, 2008.
- [11] S. Allard K. J. Levine C. Tenopir, "Design engineers and technical professionals at work: Observing information usage in the workplace," *Journal of the American Society for Information Science and Technology*, Vol. 60, No. 3, pp. 443-454, 2009.
- [12] M. Park, "User participation evaluation of a scholarly information site," *Journal of the Korean Society for Information Management*, Vol. 28, No. 4, pp. 85-97, 2011.
- [13] M. Park, "Usability of the national science and technology information system," *Journal of the Korean Bibliography Society for Library and Information Science*, Vol. 22, No. 4, pp. 5-19, 2011.
- [14] M. Park, "User-oriented Information System: Focusing on STEM Field," *Journal of the Convergence on Culture Technology*, Vol. 6, No. 1, pp. 249-253, 2020.
- [15] P. J. Agerfalk, J. Sjoström, E. Eliasson, S. Cronholm, and G. Goldkuhl, "Setting the scene for actability evaluation: understanding information systems in context," in *Proceedings of ECITE Conference*, 2002.
- [16] H. Beyer and K. Holtzblatt, *Contextual design: defining customer-centered systems*, San Diego: Academic Press, 1998.

- [17]S. Cronholm and G. Goldkuhl, "Actable information systems - quality ideals put into practice," in Proceedings of the Eleventh Conference on Information Systems, 12-14 September, 2002.
- [18]KORDI, <https://www.kiost.ac.kr>
- [19]KAERI. <https://www.kaeri.re.kr>
- [20]KOSFIC, <http://kosfic.chonnam.ac.kr>
- [21]PubMed. <https://pubmed.ncbi.nlm.nih.gov/>
- [22]BioMed Central. <https://www.biomedcentral.com>