

정보 전달 효율 향상을 위한 학술정보 시각화에 관한 연구

Study on Scholarly Information Visualization for Improving Communication Efficiency

서태설, 조성남
한국과학기술정보연구원

Seo Tae-sul, Cho Seong-nam
Korea Institute of Science and Technology Information

요약

정보의 홍수 시대에 사는 인터넷 이용자들은 많은 정보 중에서 자기가 원하는 정보만을 빠르게 찾아내고자 한다. 학술지의 경우는 대부분 문자로 서비스되고 있는데, 이를 시각화하여 서비스한다면 보다 효율적으로 정보가 전달될 수 있을 것이다. 최근 학술지 논문 기사 원문 전체가 XML로 구축되는 양이 늘어나고 있어서, XML 데이터를 기반으로 한 학술정보의 시각화 서비스에 대한 관심이 모아지고 있다. 본 연구에서는 학술정보를 시각화하기 위한 데이터 모델을 제시하고 학술정보 시각화 서비스 방안을 마련한다.

I. 서론

과학기술의 지식화 매체로서 가장 대표적인 것이 학술지라 할 수 있다. 전 세계적으로 수많은 학술지가 시시각각으로 발간되고 있고 우리나라에도 몇 천 종의 학술지가 활발하게 발간되고 있다. 하지만, 대부분의 정보가 문자 위주로 서비스되기 때문에 검색 결과를 하나하나 확인을 해야 하고, 많은 시간이 소요된다. 만일 이러한 정보를 시각화하여 서비스한다면 원하는 정보를 찾는 데 더 효과적일 수 있다^{1, 4}.

최근, 여러 분야에서 데이터나 정보의 시각화가 큰 관심을 모으고 있어서 다양한 시각화 방법과 도구들이 등장하고 있다. 본 연구에서는 학술정보를 시각화하기 위한 데이터 모델을 제시하였고, 이를 기반으로 학술정보 시각화 서비스 방안을 마련하였다.

II. 관련 연구

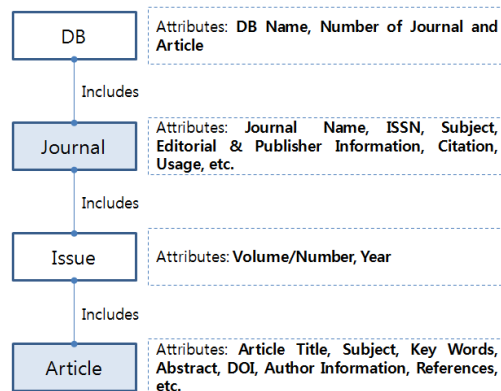
최근 빅데이터 분석 결과를 시각화하는 연구는 많으나, 학술정보를 시각화 하는 연구는 찾아보기 힘들다. 더욱이, 학술정보 시각화 연구는 저자 네트워크, 인용 관계 분석, 연구자 분포 등에 국한되어 시도되어 왔으나, 학술지와 학술지 논문 기사 단위의 시각화에 대한 시도는 많지 않았다.

학술지 논문 기사 단위의 시각화 시도는 XML(eXtensible Markup Language)로 구축되기 시작한 이후에 시도되고 있다. Shontton 등^{2, 3}은 데이터 처리 기술, 시각화 기술, 실시간 데이터 상호작용 등을 활용하여 XML로 구축된 학술지 논문 기사의 다양한 시각화를 시도하여 시멘틱 출판(Semantic publishing)의 가능성을 보였다. Seo & Choi⁴는 학술지와 학술지 논문 기사에 대

한 시각화 기초연구를 통해 데이터 모델과 서비스 전략을 제시하였다.

III. 학술정보 데이터 모델

학술정보는 일반적으로 전체 데이터(DB), 학술지(Journal), 이슈(Issue), 기사(Article)의 객체들(Object Classes)로 구성된다.

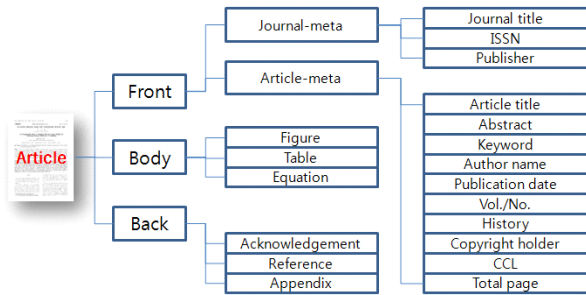


▶▶ 그림 1. 학술정보 전체 모델

그림 1에서 보는 것처럼 전체 데이터와 이슈는 정보 항목(attribute)이 단순하기 때문에 데이터 모델이 간단하나, 학술지와 기사의 경우는 많은 정보 항목을 가지기 때문에 데이터 모델을 어떻게 구성하느냐에 따라서 정보의 표현이 달라질 수 있다.

학술지 모델은 학술지를 보는 관점에 따라서 독자(reader), 투고자(author), 도서관(library), 출판사(publisher) 등으로 구분될 수 있으므로, 각각의 관점에서 필요한 항목을 선택하여 데이터 모델을 구성하면 된다.

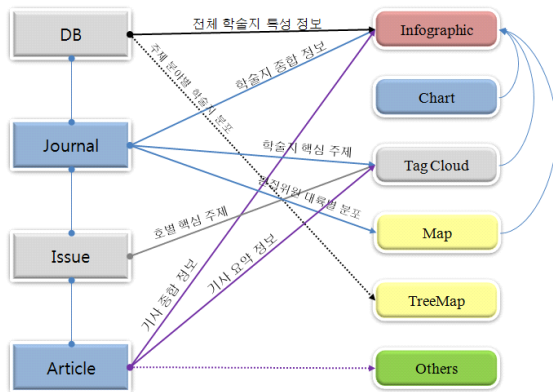
학술지 논문 기사에 대한 데이터 모델은 미국 의학도서관에서 제안한 JATS(Journal Article Tag Suite)가 대표적이다. 이것은 학술지 논문 기사 원문을 위한 XML DTD로서 NISO 표준으로 채택된 바 있다[5]. 그런데, JATS는 너무 세분화 되어 있어서 시각화에 필요한 항목만을 선별할 필요가 있다. 그림 2는 JATS의 데이터 구조를 요약한 것이다.



▶▶ 그림 2. 학술지 논문 기사 데이터 모델

IV. 학술정보 시각화 서비스 방안

학술정보는 객체별 정보 항목에 따라서 다른 시각화 방법이 적용될 수 있다. 학술정보에 적용될 수 있는 시각화 방법으로는 Infographic, Chart, Tag Cloud, Map, TreeMap 등이 있을 수 있다. 그림 3에 각 객체별로 가장 적합한 시각화 방법을 나타내었다.

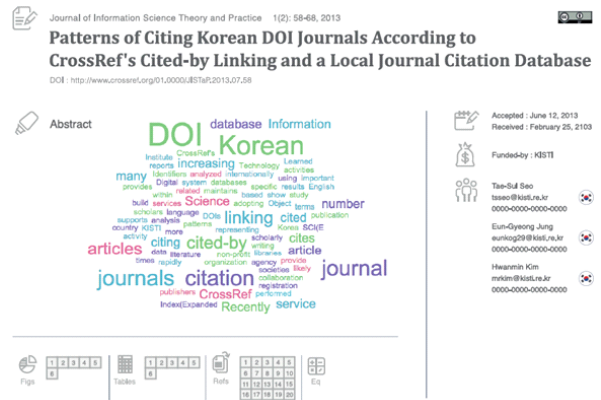


▶▶ 그림 3. 학술정보 시각화 시나리오

먼저, 전체 DB의 경우는 전체 학술지 정보를 Infographic으로 나타낼 수 있고, 주제 분야별 학술지 분포를 TreeMap으로 나타내는 것이 가능하다. 학술지의 경우는 학술지 종합 정보를 Infographic으로, 학술지 핵심 주제를 Tag Cloud로, 편집위원 대륙 분포를 세계 지도로 나타내면 된다. 이슈는 이슈 전체의 주제를 Tag Cloud로 나타내는 것 외에는 시각화가 어렵다. 마지막으로 논문 기사는 기사 전체 내용을 Infographic으로, 기사 요약 정보를 Tag Cloud로 나타낼 수 있다.

본 연구에서는 학술지와 논문 기사에 대한 시각화 서비스 화면을 설계하였다. 그림 4는 JATS 데이터 모델에

근거하여 설계된 논문 기사에 대한 시각화 서비스 설계 화면이다. 기존의 초록 대신 Tag Cloud로 직관적 개요 파악이 가능해졌으며, 그림, 표, 참고문헌, 수식의 양도 직관적으로 파악이 가능하다. 저자의 경우는 국기 이미지를 통해 출신 국가를 빠르게 알 수 있다.



▶▶ 그림 4. Infographic을 통한 학술지 논문 기사 서비스 화면

IV. 결론 및 향후 연구

본 연구에서는 학술정보 시각화 서비스를 위한 데이터 모델과 서비스 방안을 마련하였다. 학술지 데이터 모델은 이용자 요구 분석을 통해 얻었고, 논문 기사의 데이터 모델은 JATS 항목에서 취사선택하였다. 학술정보 각 객체에 대한 시각화 서비스 방법을 제시하였고, 학술지와 논문 기사에 대한 시각화 설계를 Infographic 기법으로 수행하였다.

향후 연구로는 논문 기사에 대한 Semantic Publishing 연구가 필요할 것으로 사료된다.

■ 참고 문헌 ■

- [1] 서태실 외, 학술정보 시각화 서비스 전략, pp. 108, 한국과학기술정보연구원, 서울, 2014.
- [2] Shotton, D., Portwin, K., Klyne, G., Miles, A., "Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article," PLoS Comput. Bio., Vol. 5, No. 4, pp. 1-17, 2009
- [3] Shotton, D., "Semantic publishing: the coming revolution in scientific journal publishing," Learn. Publ. Vol. 22, No. 2, pp. 85-94, 2009
- [4] Patterson, R. E., Blaha, L. M., Grinstein, G. G., Liggett, K. K., Kaveney, D. e., Sheldon, K. C. Havig, P. R., Moore, J. A., "A Human Cognition Framework for Information Visualization," Compt. Graph, Vol. 42, pp. 42-58, 2014
- [5] ANSI/NISO Z39.96-2012 JATS: Journal Article Tag Suite Version 1.0), National Information Standards Organization (NISO)