

인쇄문헌과 디지털문헌의 상호운용 연구에 관한 최근 동향

The research trend of interoperability between paper and digital documents

김지원, 숙명여자대학교 대학원

문헌정보학과

Kim, Jee-Won

Dept. of Lib. & Inf. Sci.

Graduate school at

Sookmyung Women's Univ.

사공철, 숙명여자대학교 정보과학부

Sagong, Chul

Div. of Information Science

Sookmyung Women's Univ

미래에는 종이없는 사회 (paperless society)가 도래할 것이라는 예측에도 불구하고 여전히 인쇄문헌을 선호하고 있는데 이것은 디지털문헌의 유용성에 대한 어떤 의문점을 갖게 한다. 따라서 최근 미국을 중심으로 하여 활발하게 연구되고 있는 인쇄문헌과 디지털문헌의 상호운용에 관한 동향을 개관하고 향후 이에 관한 연구과제를 제시하였다.

1. 서론

인류의 삶에 있어 커뮤니케이션의 수단은 발화되는 정보뿐만 아니라 종이에 인쇄되어진 문헌을 포함한다. 따라서 인류의 역사에 있어 도서관이란 주로 '도서'라는 인쇄매체를 통해 보존되고 계승되어진 인류의 문화·지식유산의 성스러운 보존소이었으며, 또한 인간이 보유한 정보환경에서 중심적인 역할을 해왔다. 그러나 정보통신 기술 및 컴퓨터의 발달로 인해 이러한 정보환경에 커다란 변화가 일기 시작했다. 즉 인쇄매체 중심에서 전자매체 중심의 정보환경으로의 전이가 시작된 것이다. 이와 함께 디지털도서관의 개념이 자연스럽게 대두되었다.

또한 디지털도서관 연구영역의 핵심분야 중

하나를 바로 문헌에 관한 연구로서 인쇄문헌의 디지털화 및 디지털문헌의 생성, 검색, 교환 및 관리에 대한 연구가 활발히 진행되고 있다. 특히 SGML이나 XML을 사용하여 문헌의 구조를 기술하는 방식, 즉 구조화된 문헌에 대한 연구는 디지털도서관의 등장과 더불어 각광을 받았다.

그러나 디지털도서관을 구성할 디지털문헌을 준비하는 과정에 있어 나타나는 인쇄문헌과 디지털문헌의 차이는 이용자에게 대한 정보 서비스를 원칙으로 하는 도서관의 존립을 위태롭게 하고 있다. 따라서 본 고에서는 인쇄문헌과 디지털문헌의 상호운용 연구에 관한 최근 동향을 살펴보고, 이에 관한 향후 연구과제를 제시하고자 한다.

2. 문헌의 상호운용 등장배경

컴퓨터 및 정보통신 기술의 발달은 우리로 하여금 미래에는 종이가 없는 사회 (paperless society)가 도래할 것이라는 예측을 하게 하였다. 그러나 지금은 이러한 기술이 발달하기 전보다 훨씬 더 많은 양의 종이를 소비하고 있다. 미국의 한 조사기관인 Dataquest에 따르면 1997년 미국에서는 복사기나 프린터 등을 통해 수 조(兆) 장에 이르는 양의 종이가 생산, 소비되었다고 한다 (Schilit et al. 1999. 66). 또한 우리는 매일 인터넷에 접속하며 웹을 통해 전자형태로 된 정보를 많이 얻고 있다. 그러나 오락이나 여가의 목적이 아닌 전문적 혹은 학술상의 목적을 가지고 웹 상의 문헌을 이용할 때에는 온라인의 상태에서 읽기보다는 그것을 다운로드하여 출력한 인쇄상태의 문헌으로 재조작하여 사용하는 경향이 있다. 이는 사람들이 여전히 인쇄문헌을 선호하고 있음을 보여주는 명확한 예이다.

디지털기술을 통한 문헌의 디지털화 및 디지털문헌의 생산이 진행되고 있는 이 시점에서 이러한 현상이 왜 생기는 것인가? 그 이유는 우리가 문헌에 대한 깊은 이해없이, 그리고 문헌과 인간사이에서 발생하는 다양한 상호작용의 측면에 대한 충분한 이해없이 고도로 발달된 기술을 앞세워 매체의 진이에만 급급했기 때문이다.

인쇄문헌에서 얻을 수 있는 정보는 디지털화된 문헌에서 역시 얻을 수 있어야 하고 인간과 인쇄문헌 사이에서 일어나는 다양한 상호작용이 디지털화된 문헌에서 역시 지원되어야 하는데, 지금까지의 문헌의 디지털화는 인쇄문헌이 지닌 여러 특성이나 이용자의 다양성, 그 둘간의 관계를 무시하였다. 따라서 소위 살아있는 인쇄문헌 (paper live document)이 죽어있는 디지털문헌 (digitalized dead document)으로 전이되는 결과를 낳았다. 그로 인해 인쇄문헌이 매체만 바뀌어진 디지털문헌으로 변환되는

것이 아니라 또 다른 하나의 문헌이 양산되어 고립된 정보를 생성하고, 디지털화된 문헌을 생산하는데 들인 엄청난 노력과 비용에도 불구하고 여전히 인쇄문헌을 선호하는 결과를 초래하였다.

따라서 인쇄문헌과 디지털문헌간에 존재하는 차이를 줄여주는 혹은 그 두 가지 형태의 문헌에게 가교를 제공하는 방법으로서 문헌의 상호운용에 대한 연구가 제기되었다. 문헌의 상호운용이란 인쇄문헌이 가지고 있는 다양한 정보의 손실없이 문헌의 디지털화가 이루어지게 하는 것을 의미하며, 그 목표는 인쇄문헌과 디지털문헌간의 매체 분열적인 전이가 아니라 협력적이고 통합적인 전이 (collaborative and integrated transience)가 이루어지도록 하는 것이다. 또한 이는 인쇄문헌에 있는 원 정보를 단순히 디지털 버전으로 바꾸는 것에 그치는 것이 아니라 정보의 무손실을 기본 원칙으로 하되, 거기에 이용자에게 이익을 줄 수 있는 방향으로 어떤 가치나 구조를 부여받은 지능형 문헌 (intelligent document)을 지향한다.

3. 인쇄문헌과 디지털문헌

3.1. 문헌의 개념

문헌은 오랜 세월동안 정보 및 지식관리 (information and knowledge management)에 있어 핵심적인 역할을 해왔으며, 인류의 정보 및 지식관리는 문헌발생 시점에 따라 두 가지 국면으로 나눌 수 있다. 첫째는 문헌발생 이전으로서, 이 시기는 음성문화 중심으로 지식을 표현, 관리, 전달하였으며 막대기, 소리, 춤, 의식과 같은 수단에 의존하였다. 둘째는 문헌발생 이후로서, 이 시기에는 정보 및 지식을 관리하는데 있어 문헌이라는 효과적인 수단이 등장하여, 문헌을 통한 커뮤니케이션이 발달하였다. 또한 문헌의 등장은 인류의 문화적 경제적 삶에 커다란 효과를 과급하였다 (Renear 1997. [Online]).

우리는 흔히 문헌을 정보의 집합, 또는 텍스트 정보 및 비텍스트 정보의 결합체 정도로 이해한다. 그러나 문헌은 단순한 정보의 집합 그 이상의 무엇을 가지고 있지 않다면 인류의 삶 속에서 그 오랜 세월동안 존재하지 않았을 것이다.

Levy (1994)는 문헌을 가공물 (artifact), 관련기술 (technology), 작업 (work), 장르 (genre)의 4가지 관점으로서 이해하고자 하였다.

- 가공물 (artifacts): 문헌은 물리적이고 사회적인 가공물이다.
- 관련기술 (technology): 이는 연필, 종이, 타자기, 프린터, 컴퓨터, 소프트웨어, 네트워크 등 문헌을 생성, 조작, 배포하는 것 등과 관련된 모든 기술을 일컫는다.
- 작업 (work): 문헌은 사회적 가공물이므로 궁극적인 목적은 이용이다. 따라서 문헌은 인간의 작업습관이라든가 이용행태와 관련 있으며 또한 그러한 프로세스에 포함된다.
- 장르 (genre): 앞서 언급한 세 가지 관점을 통합하는 것으로서 결국 문헌이란 분리된 가공물이 아니라 누구나 인식 또는 이해할 수 있는 유형 혹은 장르의 실체라는 의미이다. 예를 들면, 소설, 신문기사, 쇼핑목록, 리스트, 학술지 등이 장르에 포함된다. 따라서 장르는 특정한 기능이나 역할을 갖게 되며 우리는 이러한 장르로부터 문헌의 의도된 목적이나 역할을 이해할 수 있게 된다.

이와 같이 문헌은 단순히 정보만을 포함한 물리적인 객체가 아니라 사회적인 역할과 기능을 가지고 있고 이용을 전제로 하기 때문에 이용자의 작업습관에 따라 이용행태가 달라지며, 인간과의 상호작용에 따라 문헌과 연관된 활동을 야기하는 능력을 보유하고 있다.

3.2. 인쇄문헌과 디지털 문헌의 특성

인쇄문헌은 종이와 문자를 사용하여 저자의 의도를 기술한 것으로 고정성 (fixity)을 그 특징으로 가지고 있다. 또한 문헌의 크기나 외형으로 중요도를 구분할 수 있고 출판 및 배포에 비용이 많이 들며 갱신주기가 길기 때문에 최신성이 부족하다. 다양한 정보형태를 수록하기 보다는 텍스트위주의 정보를 포함하며 선형적인 검색만이 가능하다.

이와 반면에 디지털 문헌은 컴퓨터 코드, 비트맵으로 표현되는 이미지 등의 컴퓨터 기술을 이용하여 생성된 것으로 유동성 (fluidity)을 그 특징으로 가지고 있다. 디지털 문헌은 인쇄문헌과 다른 물리적인 외형을 가지고 있으며, 그 크기가 매우 다양하고 멀티미디어 정보와 같이 포함된 정보가 텍스트에 그치는 것이 아니라 사운드, 그래픽 애니메이션, 비디오 등으로 확대된다. 또한 즉각적인 수정이 가능함에 따라 문헌의 유통과 생명주기에 변화를 가져왔으며 하이퍼텍스트의 경우에서처럼 비연속적인 방법으로 문헌에 접근할 수 있다. 그리고 인쇄문헌과는 달리 디지털 문헌은 직접적인 물리적 접근을 요구하는 것이 아니라 네트워크를 통한 정보의 전달이라는 논리적인 접근을 요구한다.

<표 1> 매체에 따른 문헌의 특징 분류 (Levy 1994, 25)

인쇄문헌	디지털문헌
안정적 (stable)	불안정적 (unstable)
영구적 (permanent)	가변적 (changeable)
정적 (static)	역동적 (dynamic)
비활동적 (inactive)	활동적, 상호작용적 (active, interactive)
고정적이며 구조가 엄격함 (fixed, rigid)	유동적이며 구조가 순응적 (fluid, malleable)

4. 연구동향

4.1. 인간과 문헌간의 상호작용으로서의 문헌연관 활동에 대한 연구

사람들은 다양한 목적과 방법을 가지고 문헌을 이용한다. 그러한 목적과 방법에 따라서 사람들은 문헌을 이용할 때 문헌과 연관된 활동을 하게 된다. 이러한 활동은 인간과 문헌간의 상호작용으로 간주될 수 있으며 이는 정보읽기와 정보기록의 두 가지 형태로 나눌 수 있다. 정보읽기는 우리가 흔히 독서라고 일컫는 것이며, 정보기록은 정보읽기를 통해 정보를 추출하고 기록하는 행위로서 노트필기, 주석(comment)달기, 원문복사하기 등을 그 예로 들 수 있다.

특히 정보읽기는 문헌이용과 직접 관련이 있는데, Schilit, Price and Golovchinsky (1999)는 정보읽기 행위를 대상문헌과 연관 활동의 특성에 따라 <표2>, <표3>과 같이 분류하였다.

<표 2> 정보읽기의 상황에 따른 분류

	수동적 읽기	능동적 읽기
단일 텍스트	소설 또는 시 등을 읽는 것	교재 공부, 연구제안서 검토, 매뉴얼 검토 및 파악
다중 텍스트	전자우편을 체크, 신문을 훑어읽기, 인터넷 탐색	문제를 연구하거나 특정 분야를 조사하는 것, 전문적 지식을 계속 갱신하는 것

<표 3> 정보읽기와 연관된 활동

	수동적 읽기	능동적 읽기
단일 텍스트	텍스트에 대한 이해	텍스트 이해, 텍스트내 정보 찾기, 정보 요약하기 등
다중 텍스트	텍스트에 대한 이해, 여러 문헌 탐색, 관심있는 문헌 찾기	텍스트 이해, 여러 문헌 운반 및 탐색, 적합한 문헌 찾기, 정보에 대한 구조화

또한 O'hara et al. (1998)은 캠브리지 대학(Cambridge University)에 다니는 25명의 박사과정 학생들을 대상으로 학술정보에 대한 문헌연관 활동이 어떻게 이루어지고 있는지를 조사하였다. 조사대상 학생들이 수행한 연구활동의 여러 단계에서 발생한 문헌연관 활동을 파악하고 이에 대한 모델을 도출하였는데, 그에 의하면 학술연구 활동은 탐색, 정보검색, 정보읽기, 정보추출, 정보기록, 정보검토, 학술논문 작성의 반복적인 프로세스로 이루어진다.

Dillon (1994)은 이용자와 인쇄문헌 및 전자문헌의 상호작용에 대한 연구를 통해 TMS framework를 제시하였는데, 이는 문헌연관 활동에 영향을 주는 4가지 요소로서 TM (Task Model), IM (Information Model), MSF (set of Manipulation Skills and Facilities), SRP (Serial Reading Processor)를 의미한다.

4.2. 문헌이용 행태에 대한 연구

Bishop (1998)은 일리노이 대학의 디지털도서관 프로젝트를 통해 구축된 테스트베드의 이용자를 대상으로 구조화된 문헌의 실제 이용행태 및 이용자 연구를 실시하였다. 그는 문헌의 구성요소들이 이용자들의 연구활동이나 작업에 실제 어떻게 쓰이는지를 조사하여, 이용자의 일정한 유형이 있음을 발견하였다.

Dillon (1994)은 학술지 이용의 이유, 이용자들이 학술지에서 찾고자 하는 정보의 유형, 학술지 이용방법을 연구하였다. 특히 그는 학술지 이용자들이 정보구조에 대해 대부분 인식하고 있으며 이를 문헌구성의 원칙으로 간주하고 있다고 주장하였다. 따라서 정보구조에 대한 인식이 문헌의 적합성 정도를 판단하는데 영향을 미친다고 하였으며, 정보구조에 대한 인식 때문에 사람들이 일정한 형태를 가지고 학술지를 이용하고 있다고 주장하였다.

이외에도 O'hara and Sellen (1997)은 이용자가 온라인 문헌과 인쇄문헌을 읽는 동안에

보이는 여러 행태의 차이를 비교, 연구한 바 있다.

4.3 문헌 구조화에 대한 연구

문헌은 물리적 구조와 논리적 구조로 이루어져 있으며 이들을 여러 엘리먼트로 구조화하는 접근법은 SGML, XML 등의 마크업 언어를 통해 이미 널리 알려져 있으며 구조화된 문헌의 장점 역시 널리 인식되고 있다. 현재 많은 디지털문헌들이 혹은 디지털화된 문헌들이 이러한 접근법에 의하여 생성되었다.

Furuta (1994)에 의하면 디지털도서관의 구성요소로서 디지털문헌은 객체 엘리먼트들로 이루어진 구조화된 문헌이며 그들의 구조와 상관관계를 정의, 분석함으로써 디지털도서관을 위한 더욱 강력한 metaphor를 개발할 수 있다고 하였다. Summers (1995)는 문헌의 논리구조의 특성을 파악하여 문헌의 논리구조를 자동적으로 발견할 수 있게 함으로써 강력하면서 융통성있는 문헌조작 도구를 개발할 수 있다고 하였다.

이 밖에도 마크업 언어를 사용하여 문헌을 구조화시키고 다양한 문헌유형에 이를 적용시켜 디지털문헌 컬렉션을 구축하고자 하는 많은 프로젝트들이 진행되고 있다.

5. 지능형 문헌 (intelligent document)

“intelligent”라는 말은 본래 라틴어에서 기원한 것으로, “사이”를 의미하는 “inter”와 “선택하다, 읽다, 모으다”를 의미하는 “legere”의 결합어로서 “행간을 읽을 수 있는 존재, 혹은 지능, 판단력, 이해력을 보유한 것”을 의미한다 (WWWebster Dictionary 1999). 지능형 문헌은 인쇄문헌의 원 내용을 포함하고 있으면서, 그 전자적 형태에 구조적이면서 풍부한 표현력을 제공하는 요소나 구조를 부가하여 이용자로 하여금 검색 및 자유로운 항해 (navigation)를 할 수 있게 하는 문헌이다. 비단 이용자의 측

면에서뿐만 아니라 생산과 유통에 있어서도 쉽게 이식 (portable)할 수 있고 유지가능하며 통합할 수 있어 비용 효과적인 가치를 지닌 문헌이다.

지능형 문헌은 구조화된 정보를 바탕으로 하여 전자형태의 목차나 색인, 북마크, 간단하게 필기할 수 있는 노트공간의 제공, 구조적 정보의 제공 및 직접조작, 정보에의 효과적인 탐색이나 검색을 허용, 정보공간을 통한 비선형적인 항해 (non-linear navigation)를 허용하며, 이용자의 정보요구 및 그 잠재적 측면까지도 수용하는 정보객체가 될 것이다.

예측컨대 앞으로의 디지털문헌은 지능형 문헌을 지향할 것이며, 이는 인쇄문헌과 디지털 문헌사이의 차이를 줄일 수 있을 것이다. 또한 새로운 패러다임인 디지털 시대로의 진입을 유연하게 도와줄 방법이 될 것으로 기대한다.

6. 향후 연구과제

디지털도서관에 대한 연구 및 관련 정보기술의 개발이 매우 활성화되어 있음에도 불구하고 디지털도서관의 유용성을 높일 수 있는 요인 중 하나가 콘텐츠라는 점을 충분히 인식하지 않고 있다. 디지털도서관이 새로운 환경과 기술을 기반으로 구축되고 또한 디지털문헌으로 구성된다면 문헌의 특성과 구조를 중심으로 하는 문헌연구는 디지털도서관 연구에 있어 핵심적이라 할 수 있다.

지금까지의 문헌의 디지털화는 인쇄문헌 및 인간과의 상호작용에 대한 고려없이 개발자 혹은 관리자 중심적인 시각을 가지고 진행되었다. 그러다 보니 인쇄문헌과 디지털문헌간의 차이가 생기게 되었다. 본 고에서는 이러한 문제점을 제기하고 관련 연구가 외국에서 어떻게 전개되고 있는지 간단히 살펴보았다.

이미 미국에서는 인쇄문헌과 유사한 효과를 주는 전자책 (reading appliance)을 개발하였다. 그러나 현재 이런 분야의 연구가 국내에서는

부족한 실정이다. 따라서 향후, 인쇄문헌과 디지털문헌간의 상호운용의 측면을 더욱 깊이 있게 다루고, 인쇄문헌의 고유한 특성 및 정보유형, 관련된 인적 요소 (human factor) 등 여러 요소에 대한 연구가 필요하다.

참고문헌

- Bishop, A. 1998. "Digital Libraries and Knowledge Disaggregation: the Use of Journal Article Components", *Proceedings of the 3rd ACM Conference on Digital Libraries*. Pittsburgh, PA: ACM, p29-39.
- Dillon, A. 1994. *Designing usable electronic text: Ergonomic aspects of human information usage*. Bristol, PA: Taylor & Francis.
- Furuta, R. 1994. "Defining and Using structure in Digital Documents", *Proceedings of Digital Libraries '94*. College Station, TX: ACM, p139-145.
- Hearst, M. 1996. "Research in Support of Digital Libraries at Xerox PARC: the Changing Social Roles of Documents", *D-Lib Magazine* [online]. <<http://www.dlib.org/dlib/may96/05hearst.htm>>
- Levy, D. 1994. "Fixed or Fluid? Document Stability and New Media", *Proceedings of the European Conference on Hypertext Technology '94*. Edinburgh, Scotland: ACM, p24-31.
- Levy, D. and Marshall, C. 1995. "Going Digital: a Look at Assumptions Underlying Digital Libraries", *Communications of the ACM*, 38(4): 77-84.
- O'hara, K. et al. 1998. "Student Readers' Use of Library Documents: Implication for Library Technology", *Proceedings of CHI '98*, LA, CA: ACM, p233-240.
- O'hara, K and A. Sellen. 1997. "A Comparison of Reading Paper and Online Documents", *Proceedings of CHI '97, Conference on Human Factors in Computing Systems*. Atlanta, GA: ACM, p335-342.
- Renear, A. 1997. "The Digital Library Agenda", *D-Lib Magazine* [online]. <<http://www.dlib-org/dlib/july97/07renear.htm>>
- Schilit, B. et al. 1998. "Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations". *Proceedings of CHI '98*. LA, CA: ACM, p249-256.
- Schilit, B. et al. 1999. "As We May Read: the Reading Appliance Revolution". *Computer* 32(1), IEEE : 65-73
- Summers, K. 1995. "Toward a Taxonomy of Logical Document Structures". *Proceedings of DAGS '95: Electronic Publishing and the Information Superhighway*. Boston, MA[online]. <<http://www.cs.darmouth.edu/~samr/DAGS95/Papers/summers.htm>>
- Webster Dictionary*. 1999. [online].
- Merriam-Webster, Inc., [cited 1999.7.15]. <<http://www.m-w.com/dictionary.htm>>