

# 디지털 콘텐츠의 어노테이션에 관한 연구

곽 승 진

충남대학교 문헌정보학과 교수

[sjkwak@cnu.ac.kr](mailto:sjkwak@cnu.ac.kr)

# 목차

1. 서 론
2. 어노테이션의 개념
  - 2.1 어노테이션의 정의
  - 2.2 어노테이션의 필요성
  - 2.3 어노테이션 시스템의 구조
3. 어노테이션 시스템 사례연구
  - 3.1 Annotea 시스템
  - 3.2 DiLAS 시스템
4. 어노테이션의 메타데이터 형식
  - 4.1 NSDL 어노테이션
  - 4.2 DLESE 디지털도서관 어노테이션
5. 어노테이션을 이용한 정보서비스 개발
  - 5.1 웹 슈퍼구조
  - 5.2 디지털도서관과 어노테이션
  - 5.3 전자책, 전자저널과 어노테이션
6. 결론 및 제언

## 1. 서론

정보의 홍수시대인 오늘날을 살고 있는 현대인들은 필요한 정보를 선택하고 가치가 없거나 불필요한 정보를 제거하는데 많은 시간과 노력을 소비하지만 종종 본인의 정보 선택에 대하여 불만족해 한다. 따라서 디지털 정보 환경에서 보다 효율적인 정보접근과 선택을 위한 도구가 필요하게 되었고, 이러한 도구를 구축하는 진보된 기술 중의 하나가 메타데이터(metadata)와 어노테이션(annotation)이다.

인터넷 붐 이전에도 인쇄매체와 시청각 자료에서 이용 가능한 정보는 범람하였으나, 기존의 정보매체와 인터넷의 정보 이용에 차이점은 분명하다. 인쇄매체는 사용자들이 필요한 정보를 선택하기 위한 일반적인 프레임워크가 구축되어 있다. 대표적으로 초록, 서평, 평판 등과 SCI(Science Citation Index), JCR(Journal Citation Reports), Books in Print 등의 서지도구들이 이용자들의 정보 선택을 도와준다. 그러나 인터넷 정보는 지능형 검색엔진들과 메타데이터를 활용하는 디지털도서관들이 개발되고 있지만 아직까지 필요한 정보만을 선택하기에는 어려움이 많다.

메타데이터는 구조화된 정보로서 자원의 소재 파악, 식별, 기술(description), 접근과 관리를 지원하여 인터넷 자원의 이용과 공유를 증진하는 효과적 도구이다. 디지털 환경에서 어노테이션은 원본문서와 문서의 특별한 부분에 부가적인 설명과 강조를 위한 표시나 어구를 추가로 기록하는 것으로 메타데이터보다 더 디지털 콘텐츠의 효율적인 검색과 이용을 위한 여러 가지 장점을 갖고 있다. 그러나 두 개념은 서로 충돌하지 않고 상호의존적이다.

보통 메타데이터를 “데이터에 대한 데이터”라고 하면 어노테이션은 “콘텐츠에 대한 콘텐츠”를 의미한다. 그래서 어노테이션 자체가 콘텐츠이고 또한 메타데이터의 기능이 있다. 어노테이션은 텍스트, 음성 설명이나 소견(remarks), 인덱스, 마크업(markups), 하이퍼링크 등을 포함하므로 메타데이터보다 범위가 더 포괄적이며 구조 또한 더 복잡하다(Nagao 2003).

어노테이션은 텍스트뿐만 아니라 이미지, 음성, 동영상 등을 포함할 수 있다는 점에서 유용하고, 이러한 콘텐츠로부터 지식의 발견이 용이하고 하이퍼텍스트의 단순성과 정적인 면을 극복할 수 있다. 또한 시각장애인에게는 문장에 대한 음성을 제공함으로써 정보를 제공하는 등 핸디캡을 가지고 있는 이용자에게 유용하기 때문에 디지털 정보환경에서 어노테이션에 대한 연구의 필요성이 증대되고 있다.

본 연구의 목적은 최근 디지털 환경에서 다양한 장점을 갖고 있는 어노테이션 시스템의 구조를 분석하고 사례조사를 통하여, 웹 서비스와 디지털도서관 그리고 전자책 등의 디지털 콘텐츠에 사용자의 다양성을 고려한 효율적인 정보 접근과 재사용(reuse), 공유를 증진하는 어노테이션의 적용 방안을 제시하는 것이다.

## 2. 어노테이션의 개념

### 2.1 어노테이션의 정의

사람은 누구나 책이나 문서를 읽을 때 중요한 부분에 강조, 해설, 설명을 하기 위해 표시를 하거나 부가적인 정보를 입력한다. 이와 같이 원본문서의 특별한 부분과 정보의 다른 부분에 추가되는 부가 정보를 어노테이션이라고 한다. 일반적으로 어노테이션은 문서나 문서의 특별한 부분에 대한 개인의 의견, 해석, 강조, 비판 등의 목적으로 생성되는 정보를 의미하며, 하이라이트, 밑줄(line), 도형(symbol), 노트(note) 등과 같은 형태로 표현된다(Marshall 1997).

전통적인 의미의 어노테이션은 문헌정보학용어사전에서 목록이나 서지에 추가되는 주기(note)이며, 도서의 주제나 내용을 설명하거나 평가, 또는 기술한 것을 뜻하며(사공철 등편 1996), 문헌이나 그 내용에 관한 간략한 비평이나 설명(Rubin 2000)이라고 정의하였다.

컴퓨터공학 분야에서 어노테이션은 프로그램 실행에는 영향을 주지 않고 사람이 정보를 명확하게 관독하는데 도움을 주기 위하여 프로그래머가 컴퓨터 프로그램, 제어언어 또는 데이터 집합 내에 삽입해 넣는 정보다. 이는 다른 프로그래머나 잠재적인 프로그램 코드 이용자에게 프로그램의 개선 등을 위하여 힌트나 계획을 설명하는 것이다. 이러한 정보는 콘텐츠의 의미를 추론하는데 도움이 된다(Wikipedia 2006).

생물학분야, 이미지처리, 언어학 등 특별한 영역에서도 어노테이션이 사용된다. 분자생물학(molecular biology)과 생명정보학(bioinformatics)에서는 1980년대부터 DNA 유전자 어노테이션이 필요하게 되었으며, 디지털 이미지분야에서는 보통 이미지에 접착성 노트(포스트잇), 원, 화살표 등을 원본 이미지의 변경 없이 시각적인 메타데이터를 첨가하는 것으로 사용된다. 언어학적(linguistics) 어노테이션은 언어적 형태의 정보를 추가한다.

### 2.2 어노테이션의 필요성

어노테이션의 생성은 기억을 돕기 위하여, 정보적 도움을 주기 위한 것, 자신의 생각을 표현하고 의견의 공유가 주된 목적이다(Denoue, Vignollet 2000). 전자문서에서 어노테이션은 종이문서와 달리 문서 내용에 대한 링크, 해석, 질문 기능뿐만 아니라 입력된 어노테이션에 대한 검색, 편집, 공유 등의 기능을 제공한다. 따라서 어노테이션은 일회성 정보가 아니고 재사용과 공유될 수 있는 중요한 정보로 웹 환경에서 더욱 그 필요성이 요구된다.

디지털 환경에서 다양한 문서와 데이터베이스 관리 응용분야에서의 웹 어노테이션은 개념은 매우 유용한 기법으로 다양한 장점이 있다(Perry 2001).

- 사용의 자취를 제공(provide a trace of use)
- 재삼자의 논평(third party commentary)
- 정보 공유(information sharing)
- 정보 필터링(information filtering)
- 문서 내용에 대한 의미적 레이블링(semantic labeling of document content)
- 검색능력의 향상(enhanced search)

웹 어노테이션은 웹 정보원과 관련된 온라인 어노테이션이다. 웹 어노테이션을 이용하는 사용자는 실제적으로 정보원 자체를 수정하는 것 없이 정보를 추가하고 변경하고 제거할 수 있다. 또한 어노테이션 시스템은 사회적 소프트웨어 도구이고 시멘틱(semantic) 웹의 기반이다. Wikipedia에서 웹 어노테이션은 다음과 같이 사용될 수 있으므로 사용성이 매우 높다고 주장하였다.<sup>1)</sup>

- 웹 정보원의 평가(유용성, 이용자 친근성, 소수 사용자를 위한 적합성 등)
- 자료의 추가나 삭제를 통한 콘텐츠의 개선이나 각색
- 협력적 도구: 특정한 정보원의 콘텐츠에 대해 토론하는 것
- 예술적 또는 사회적 비평 도구: 웹상에 나타난 기관이나 사상에 대하여 웹 사용자에게 재해석하고 풍부하게 하는 것 또는 이의를 제기하는 것을 허락하는 것

인터넷 환경에서 어노테이션의 필요성이 증대된 만큼 웹 어노테이션 도구는 이미 존재하는 웹 페이지에 개인적 또는 공개적인 어노테이션이 가능해야하고, 종이문서와 같은 방법으로 전자문서에도 노트할 수 있어야 한다. 또한 웹 어노테이션은 정보소비자로서 독자가 논평을 추가하고, 새로운 관계를 만들고, 콘텐츠를 해석하고 또한 웹의 구조와 내용의 융합을 촉진할 수 있어야 한다고 주장하였다(Perry 2001).

### 2.3 어노테이션 시스템의 구조

현재 웹 어노테이션에 관한 연구는 멀티미디어 분야 등 매우 다양한 분야에서 활발하게 진행되고 있으며 유용한 시스템들이 개발되고 있다. 어노테이션 시스템의 구조는 연구자의 관점에 따라 다양하게 설명된다. 어노테이션을 두개의 영역으로 구분하면 문서 내에서 어노테이션의 위치를 파악하기 위하여 사용되는 앵커(anchor)와 어노테이션 내용과 작성자에 관한 정보로 구성된다. 대표적인 사례로 웹 어노테이션 시스템인 Annozillas는 어노테이션 내용은 서버에 RDF로

---

1) [http://en.wikipedia.org/wiki/Web\\_annotation](http://en.wikipedia.org/wiki/Web_annotation)

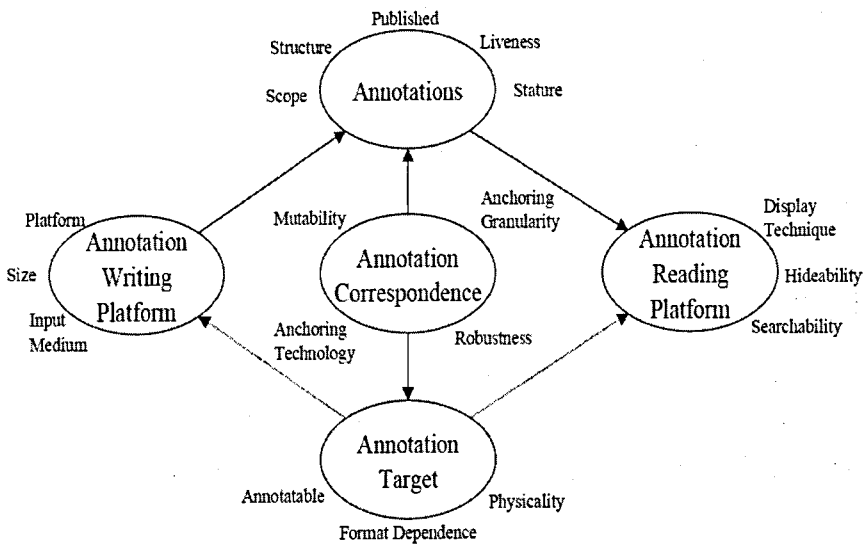
저장하고 주석되어진 웹 문서의 구역 식별자로 Xpoint를 사용한다.

Denoue와 Vignollet(2000)는 어노테이션 시스템을 크게 어노테이션의 생성, 출력, 저장 모듈로 구성하였으며 시스템 구조는 <표 1>과 같다.

<표 1> 어노테이션 시스템 구조

구분	내용
어노테이션 생성	문서 내에 키보드, 펜, 마우스 그리고 다양한 메뉴들을 사용하여 밑줄, 하이라이트, 원, 사각형 등의 여러 형태의 어노테이션을 생성
어노테이션 출력	문서 내에 생성된 다수의 앵커와 어노테이션들을 사용자들에게 제공
어노테이션 저장	문서와 문서에 입력된 어노테이션 정보를 저장. 최근에는 저작권법의 영향과 재사용을 위해 문서와 어노테이션 정보를 서로 분리하여 저장하는 방법을 많이 사용

컴퓨터 공학 관점에서 어노테이션 시스템 구조는 어노테이션의 입력 부분(annotation writing platform), 출력(annotation reading platform), 대상(annotation target), 내용(annotations), 어노테이션과 어노테이션 대상의 교신(annotation correspondence)의 다섯 개의 구성요소로 구성된다고 주장하였다(Cousins, Baldonado, and Peapcke 2000). 컴퓨터 공학 관점에서 제시된 어노테이션 시스템 구성요소는 <그림 1>과 같다.



<그림 1> 어노테이션 시스템 구성요소

Annotator 어노테이션 시스템은 어노테이션을 내용(content), 형태(form), 기능(functionality)으로 정의하였다. 형태는 밑줄, 하이라이트, 원 등이며 기능은 어노테이션이 기억, 공유, 해석 등 어떠한 목적을 위해서 사용하는지를 표현하였다. 그리고 어노테이션과 연관된 앵커 정보도 원본 문서와 별도로 저장하고, 어노테이션에 대한 주제도 지정할 수 있도록 디자인하였다(Denoue and Vignollet 2000).

어노테이션 시스템의 구성요소들은 사용자가 편리하게 어노테이션의 형식과 내용을 만들고 정보를 공유하고 의견을 교환하는 것을 효과적으로 지원하기 위하여 설계되어야 한다.

### 2.3.1 어노테이션 입력 인터페이스

어노테이션은 개인이 습득한 지식의 표현이며 책이나 문서의 중요한 부분을 강조하거나 설명할 때 사용하며, 일반적으로 밑줄, 하이라이트, 원이나 사각형 등의 도형, 포스트일과 같은 쪽지 등으로 표현된다.

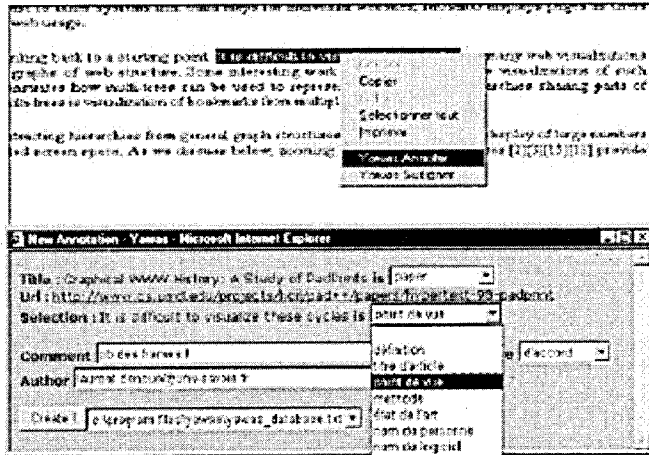
어노테이션의 생성 인터페이스는 크게 네 가지 유형으로 나눌 수 있다. 게시판과 같은 챗\_보딩 인터페이스, 라인기반 인터페이스, 심볼기반 인터페이스, 종이와 같은 유형의 인터페이스로 <표 2>와 같이 구분할 수 있다.

<표 2> 어노테이션 입력 인터페이스 유형

구분	특징
CBI (Chat-Boarding interface)	동일한 주제에 관한 다중사용자간의 의견을 쓰레드 기반 메시지로 표현(예:게시판)
LBAi (Line-Based Annotation interface)	라인, 하이라이트, 노트, WIMP(Window, Icon, Mwne, Pointingdevice or Pull-down menu)
SBAi (Symbol-Based Annotation interface)	타원, 박스, 별모양, 자유형 표시 등
PLAi (Paper-Like Annotation interface)	펜기반의 자유형 마킹, 자유로운 공간 레이아웃

매우 간단하고 편리한 프로그램으로 누구나 웹 페이지를 하이라이트 할 수 있는 어노테이션 시스템이 YAWAS(Yet Another Web Annotation System)다.<sup>2)</sup> 모든 기능은 인터넷 익스플러러를 통해서 접근할 수 있으며 쉽게 어노테이션을 반입하고 반출할 수 있다. 이 시스템은 Javascript로 개발되었으며 어노테이션을 시스템 서버에 저장하기 때문에 편리한 Google 인터페이스를 통해 쉽게 검색이 가능하다. YAWAS의 입력 인터페이스 화면은 <그림 2>와 같다.

2) <http://www.fxpal.com/people/denoue/yawas>



〈그림 2〉 YAWAS의 입력 인터페이스 화면

### 2.3.2 어노테이션 출력 인터페이스

어노테이션 시스템에서 출력 인터페이스는 사용자가 어노테이션 및 어노테이션 대상과 상호작용하게 하는 것이다. 또한 본인이 입력한 어노테이션 정보뿐만 아니라 다른 사람이 입력한 어노테이션 정보를 확인하는 방법으로 사용된다. 전자문서에서 어노테이션의 표현도 종이문서에서 사용되는 어노테이션 방법처럼 매우 다양하지만 일반적으로 완전히 다른 시스템이다.

웹 어노테이션의 표현 기술은 어노테이션 작성자와 독자가 원본 문서로부터 어노테이션 내용을 구분할 수 있도록 하는 것으로, 표현 방법을 네 가지로 구분할 수 있다. 어노테이션을 원본문서와 분리하여 표시하는 독립적(independent) 출력, 원본문서 내에 겹쳐서 표시하는 중첩(overlaid) 출력, 원본 문서의 앵커와 인접한 위치에 나란히 표시하는 인라인(inline) 출력, 그리고 애니메이션 기능을 이용하여 표시하는 유동적(fluid) 출력으로 구분된다(Cousins, Baldonado, and Peapcke 2000).

#### 1) 독립적 어노테이션 출력

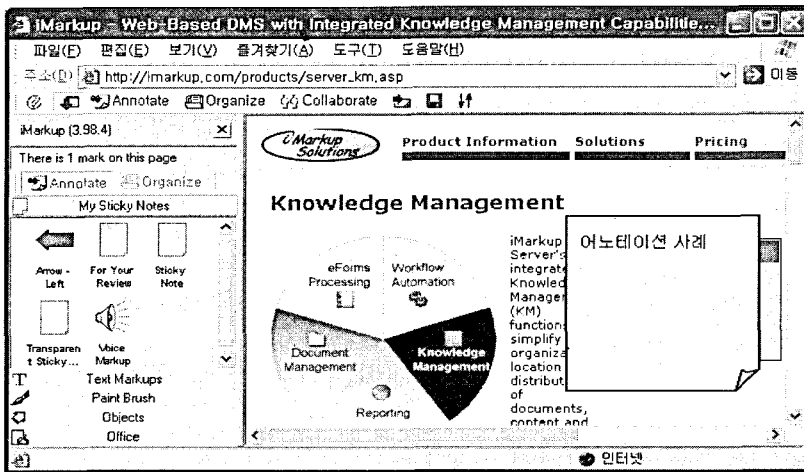
어노테이션을 원본 문서와 분리하여 별도의 윈도우에 출력하는 방법이 독립적 출력이다. 음성 어노테이션과 이메일로 전송되는 어노테이션 등이 여기에 해당되며 메타데이터가 종종 이러한 방법으로 표현된다.

독립적 어노테이션 출력은 원본문서의 변경이나 가려짐 등의 문제점을 유발하지는 않지만, 어노테이션과 앵커와의 관계를 파악하기 힘들고 자연언어처리 기법 등을 사용할 때 누락되는 단점이 있다. 독립적 어노테이션 출력을 제공하는 시스템으로 W3C에서 제안한 Annotea와 HyperNews, CoNote 등이 있다.



### 2) 중첩 어노테이션 출력

어노테이션을 원본 문서의 앵커에 가까운 곳에 내용이나 여백에 겹쳐서 표시하는 방법이 중첩 어노테이션 출력이다. 이 방법은 어노테이션의 크기와 위치가 고정적이지 않아 간단하고 강력하여 시각적 차별성이 좋다. 그러나 어노테이션과 앵커의 관계를 파악하기 어렵고 문서가 가려지는 단점이 있으며, 여러 개의 어노테이션을 표시하기 힘들다. 대표적인 기술로는 Post-IT을 적용한 iMarkup<sup>3)</sup> 시스템이 있으며 iMarkup 중첩 어노테이션 화면은 <그림 3>과 같다.



<그림 3> iMarkup 중첩 어노테이션

### 3) 인라인 어노테이션 출력

인라인 어노테이션은 원본문서에서 어노테이션과 앵커와 관계를 쉽게 파악할 수 있는 방법이며, 어노테이션을 원본문서에 통합하여 표시하는 기법으로 전통적으로 많이 사용되고 있다. 원본문서와 어노테이션을 구별하기 위하여 텍스트 크기, 폰트, 색깔 등을 표시하는 다양한 방법이 있다. 어노테이션이 원본문서에 통합되어 표시됨으로 페이지의 레이아웃 속성에 영향을 줄 수도 있다.

인라인 어노테이션 출력 시스템으로 워터루대학에서 개발한 CritLink<sup>4)</sup>와 ComMentor가 있다. CritLink 시스템에서 어노테이션은 서버에 저장되고 어노테이션은 제작과 동시에 문서에 순차적으로 생성된다. 어노테이션들은 네 가지 기호를 사용하여 의견제시(comment), 문제제기(issue), 질의(query), 동의(support)를 표시한다. CritLink는 입력 순서에 따라 어노테이션이 중첩될 수 있으며, 마우스 오버 기능을 이용하여 인라인 출력과 네 가지 기호를 사용하여 새로운 창에 상세

3) [http://www.imarkup.com/client/annotate\\_page.asp](http://www.imarkup.com/client/annotate_page.asp)

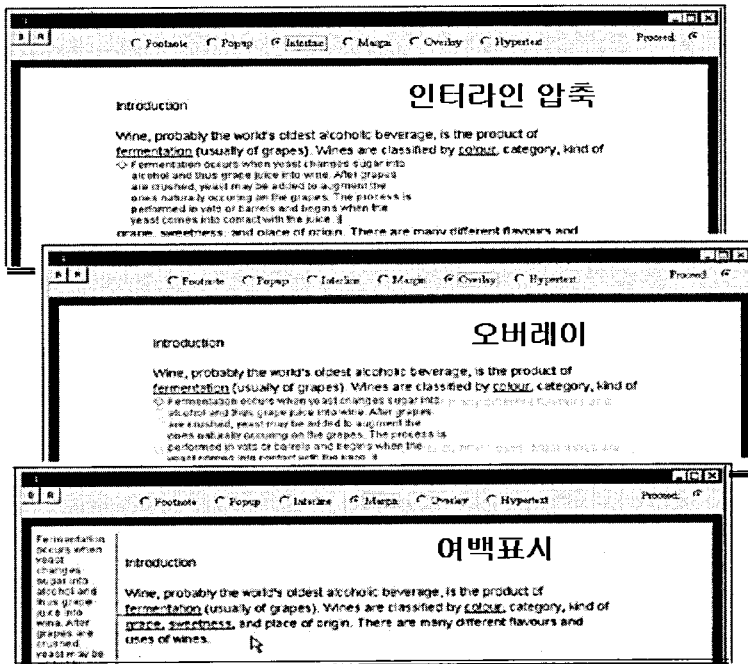
4) <http://zesty.ca/crit/>

한 어노테이션을 표시할 수 있다.

4) 유동적 어노테이션 출력

유동적 어노테이션의 출력은 애니메이션 기능을 사용하여 어노테이션을 보여주는 방법으로 인터라인(interline) 압축, 오버레이(overlay), 여백표시(margin) 방법이 있다(Zellweger et al 2001). 유동적 어노테이션 출력 인터페이스 화면은 <그림 4>와 같다.

인터라인 압축은 어노테이션을 앵커와 다음 라인 사이에 한 라인씩 점차적으로 읽을 수 있는 크기로 디스플레이 되는 방법이며, 다음 줄은 자동으로 어노테이션 라인 수 만큼 아래로 이동한다. 오버레이 방법은 원본문서를 이동시키지 않고 원본문서를 투명하게 표시한 후 그 위에 어노테이션을 겹쳐서 표현하는 방법이다. 여백표시 방법은 앵커와 가까운 곳에 여백을 만들어서 어노테이션을 나타낸다. 그러나 이 방법은 사람들이 원본문서와 어노테이션 사이를 빠르고 자연스럽게 이동할 수 있어서 편리하지만, 하나의 앵커에 다수의 어노테이션이 생성될 수 없다는 단점이 있다.



<그림 4> 유동적 어노테이션 출력 인터페이스

2.3.3 어노테이션 저장 구조

어노테이션 시스템의 저장 방법은 크게 두 가지로 구분된다. 첫 번째는 전통적인 방법으로 원

본문서와 어노테이션을 통합하여 저장하는 방법이다. 두 번째 방법은 원본문서로부터 어노테이션을 분리하여 저장하는 방법으로 최근에 개발된 대부분의 어노테이션 시스템에 적용되고 있다.

통합저장 방법은 어노테이션의 재사용을 고려하지 않으므로 분리된 관리 시스템 또한 필요하지 않다. 그러나 대부분의 어노테이션 시스템은 원본문서와 어노테이션과의 관계를 파악하기 위하여 분리된 관리 시스템이 필요하다.

매우 다양한 사용성을 고려한 어노테이션 시스템은 원본문서의 구조정보와 어노테이션 정보의 앵커를 저장해야 한다. 원본문서의 구조정보를 사용하기 위해서는 구조화된 문서를 사용하고, 어노테이션을 저장할 때 원본문서와 앵커의 구조정보를 포함하여 저장한다. 어노테이션의 저장 구조 사례는 <그림 5>와 같다.

```

<annotation>
  <element_name> // Element of the existing documents
                    including anchors
  </element_name>
  <anchor_string> // Anchor character </anchor_string>
  <date> // Date of Annotation creation </date>
  <user_id> // Annotation creator </user_id>
  <comment> // Annotation contents inputted by user</comment>
</annotation>
    
```

<그림 5> 어노테이션 저장 구조 사례

독립 어노테이션 시스템의 저장 구조는 Annotea와 iMarkup, DiLAS(Digital Library Annotation Service) 시스템 사례에서 살펴볼 수 있다. Annotea 시스템은 다양한 어노테이션의 재사용을 고려한 시스템으로 오픈 RDF를 기반으로 정의된 원본문서와 어노테이션으로써 구조화된 문서를 사용한다. Xpointer와 Xlink를 사용하여 앵커와 어노테이션을 표시하고 저장하고, 앵커정보를 확인하기 위하여 Xpointer와 Xlink를 통하여 원본문서를 참조한다(Nagao 2003).

iMarkup 시스템은 다양한 메뉴와 도구들을 사용하여 텍스트, 접착성 노트, 하이라이트, 음성 어노테이션 등을 만들어 저장한다. 그러나 원본문서로부터 어노테이션을 분리하여 저장하기 때문에 원본문서와 앵커정보의 구조에 영향을 주지 않는다.

### 2.3.4 어노테이션 인터페이스 디자인

웹 어노테이션 시스템의 문제점은 새로운 윈도우로 어노테이션 정보를 표시할 때 원본문서와 앵커의 관계가 불분명하고, 원본문서와 결합하여 표시하면 원본문서가 가려지거나 페이지 레이아웃이 변경되는 등의 문제이다. 또한 하나의 앵커에 많은 어노테이션이 생성되었을 때 유용한 어노테이션이 어떤 것인지 발견하기 어려운 문제점이 있다.

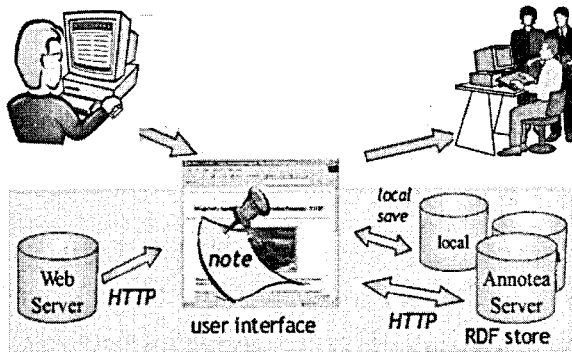
어노테이션 시스템의 인터페이스 디자인 설계 시에 다음과 같은 사항이 고려되어야 한다. 첫째, 원본문서를 읽기 쉽게 해야 한다. 앵커의 표시나 입력된 어노테이션으로 인하여 원본문서를 읽는데 어려움이 없게 디자인하고 또한 원본문서에서 어노테이션이 쉽게 구별되도록 한다 (Ovsiannikov 1999). 둘째, 효과적으로 어노테이션 정보를 제공해야 한다. 즉 최초의 접근으로 많은 어노테이션 정보들에서 더 유용한 정보를 선택할 수 있도록 해야 한다. 셋째, 앵커와 어노테이션 사이의 관계를 명확하게 표시하여야 한다. 사용자가 어노테이션 정보를 확인하려고 할 때 어떻게 앵커들과 어노테이션들이 입력되었는지 발견할 수 있어야 한다.

분산된 정보환경에서 어노테이션 시스템은 사용자 관심사를 고려한 필터링 기법과 우선 접근을 위한 순위부여 방법, 다중 사용자 환경에서의 어노테이션 정합, 유동적·계층적인 어노테이션의 시각화 기법 등을 고려한 개발이 필요하다.

### 3. 어노테이션 시스템 사례연구

#### 3.1 Annotea 시스템

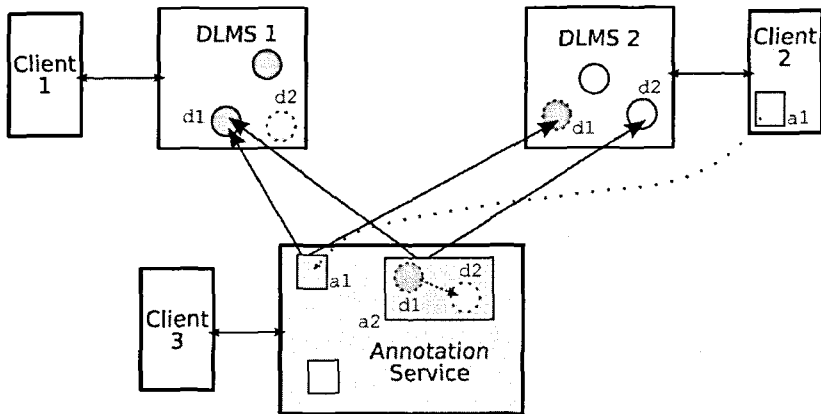
Annotea는 웹 기술의 재사용을 목적으로 RDF 구조에 기반한 웹 어노테이션 시스템이며, 메타데이터 기반의 태그, 북마크, 어노테이션의 공유를 통하여 협력을 증진한다. Annotea는 메타데이터로서 어노테이션을 기술하기 위하여 RDF 스키마를 사용하고, 주석처리된 문서에서 어노테이션의 위치를 식별하기 위하여 XPointer를 사용한다. 또한 간단한 북마크 스키마로 북마크와 토픽(topic) 메타데이터를 기술한다(Annotea Project Homepage 2006). Annotea 시스템의 기본 구조는 <그림 6>과 같다.



<그림 6> Annotea 시스템 기본 구조

### 3.2 DiLAS 시스템

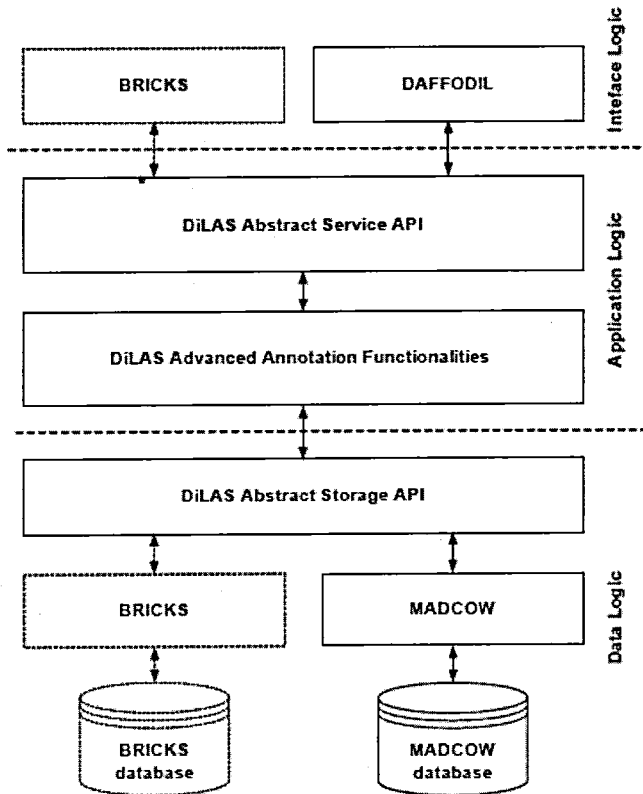
과거의 어노테이션 시스템은 문서와 함께 어노테이션을 저장하고 그것을 조회하기 위하여 중앙집중식 디지털도서관 리파지토리(repository)를 이용한다. 그러나 분산된 정보환경과 서비스 중심(service-oriented) 정보구조에서는 각각의 디지털도서관 관리 시스템(DLMS)에서 어노테이션을 독립적으로 관리하는 것이 필요하며, 이러한 목적을 달성하기 설계되고 개발된 프로젝트가 DiLAS(Digital Library Annotation Service)다. 이 시스템은 기존의 디지털도서관 시스템에서 지식의 생성과 공유, 상호작용을 촉진하기 위해 개발된 독립 어노테이션 서비스로 어노테이션을 원본문서와 분리하여 관리한다(Agosti et al, 2006). DiLAS의 독립 어노테이션 서비스 개념도는 <그림 7>과 같다.



<그림 7> DiLAS 독립 어노테이션 서비스

이런 서비스를 이용하면 개별 디지털도서관이 시스템의 기본 구조를 변경하지 않고 ASP 방식 등을 통해서 인터페이스의 변경만으로도 주석 기능을 도입할 수 있으므로 국내에도 검토할 만한 방식이라고 생각된다(이재윤, 황혜경 2006).

DiLAS는 분산 어노테이션 서비스를 지원하고 평가할 수 있는 구조로 디자인되었으며, 시스템의 구조는 데이터, 어플리케이션, 인터페이스 로직의 3계층으로 구성되어 있다. DiLAS 서버는 다른 DL시스템을 통합하여 다양한 이용자들이 새로운 어노테이션을 만들고 검색하고 활용할 수 있도록 하는 장점이 있다. DiLAS 시스템 구조(Agosti et al, 2005)는 <그림 8>과 같다.



〈그림 8〉 DiLAS 시스템 구조

DiLAS의 최초 프로토타입은 서로 다른 디지털도서관 시스템인 FAST(Flexible Annotation Service Tool), MADCOW(Multimedia Annotation of Digital content Over the Web)를 BRICKS 시스템과 통합하는 것이며, 최근에 DAFFODIL(Distributed Agents for User-Friendly Access of Digital Libraries)을 DiLAS에 통합하였다.

## 4. 어노테이션의 메타데이터 형식

### 4.1 NSDL 어노테이션

NSDL(National Science Digital Library)은 미국의 국가과학디지털도서관으로서 과학(Science), 기술(Technology), 공학(Engineering), 수학(Mathematics)분야의 교육과 연구를 지원하는 국

가 온라인 도서관이다. 이 디지털도서관은 분산된 정보환경에서 모든 단계의 교사와 학생을 위한 많은 정보원으로부터 확실하고 품질 높은 디지털자원을 접근하기 위한 디지털도서관으로 미국과 학재단(NSF)이 지원하는 방대한 프로젝트다.

NSDL 메타데이터 리파지토리는 장서(collections)와 아이템(items)의 두개의 기본 레코드 타입이 있다. 현재 어노테이션 레코드는 계획중이며 각 아이템, 장서와 관련된다. 어노테이션 레코드는 자원에 대한 설명이나 정보이며 다른 정보에 대한 링크로 구성된다. NSDL에 제안된 어노테이션 초안(draft)의 표현형식은 13개의 메타데이터 요소로 정의 되었으며, DLESE 어노테이션에 영향을 주었다. NSDL 어노테이션의 메타데이터 구조는 <표 3>과 같다.

<표 3> NSDL 어노테이션 메타데이터 구조

네임스페이스	요소명	정의
*DC	Creator	자원의 내용을 만드는 것에 책임이 있는 존재
DC	Title	자원에 대한 이름
ANNO	Created	어노테이션이 만들어진 날짜와 시간
DC	Date	가장 최근에 수정된 날짜
ANNO	Body	어노테이션의 본문
NSDL	Service	어노테이션의 출처
NSDL	Type	어노테이션의 유형(현재 시험적으로 조정된 Type 종류가 진행중)
DC	Annotates	어노테이션 자원과 어노테이션이 적용될 자원의 관계
DC	Identifier	주어진 문맥내에서 자원의 명확한 식별자
NSDL	ID	어노테이션의 식별자
NSDL	LINK_ID	어노테이션 메타데이터에 의해 참조된 NSDL MR 자원의 식별자
ANNO	Context	가장 직접적으로 적용된 어노테이션으로 자원 안의 Context
NSDL	Status	리뷰 또는 어노테이션의 상태

\* 이 제안은 표준 DC(Dublin Core)정의를 수정하여 사용함

현재 NSDL 서비스에서 시험적으로 진행중인 어노테이션 유형은 다음과 같다.

- 조언(advice): 독자에게 조언을 표시하고 있는 어노테이션의 하위분류
- 어노테이션(annotation): 어노테이션의 일반 형태를 기술하는 상위분류
- 모아진 색인의 평균 값
- 변경(change): 원본문서에 상세하게 기록하거나 변경을 신청하는 어노테이션
- 논평(comment): 논평을 주석하는 어노테이션의 하위분류
- 편집자의 요약(editor's summary)
- 예(example): 예를 표현하는 어노테이션의 하위분류
- 설명(explanation): 콘텐츠의 설명을 표시하는 어노테이션의 하위분류

- 도전적인 교육과 학습상황정보(information on challenging teaching and learning situation)
- 질문(question): 내용 대한 질문을 표시하고 있는 어노테이션의 하위분류
- 리뷰(review): 공식적 또는 비공식적 평가
- 참조(see also): 다른 자원의 참조를 표시하는 어노테이션의 하위분류
- 교육 팁(teaching tips)

NSDL 어노테이션 요소는 Annotea, DLESE 등의 특별한 요소를 반영하였으나, 실제적으로 다른 DC 요소와 한정어를 어노테이션 메타데이터 레코드도 받아들일 수 있어야 한다.

#### 4.2 DLESE 디지털도서관 어노테이션

DLESE(Digital Library for Earth System Education)는 지구시스템 교육을 위한 디지털도서관으로 지구와 관련된 교육과 학습을 위한 우수한 디지털 자원의 발견, 제작, 문서화, 배포, 이용할 수 있는 지원을 제공한다. 이 시스템은 NSDL 프로젝트의 하나로 과학과 공학분야에 포함되며, 커뮤니티 회원들의 지식과 통찰력 있는 어노테이션이 포함된 장서에서 교육 자원의 가치를 높일 수 있다.

DLESE는 교육 자원과 장서, 어노테이션을 기술하는 다른 유형의 강력한 메타데이터 구조를 사용한다. 자원 메타데이터는 유일한 식별자와 교육 자원을 기술하는 필수적인 정보를 포함하고, 장서 메타데이터는 자원의 주제, 조직, 교육전략 등을 기술한다. 어노테이션 메타데이터는 자원 레코드에서 분리된 교육 자원에 대한 추가적인 정보를 포함한다.

Kastens(2005)는 DLESE 어노테이션을 역사적 관점(Historical Perspective) 어노테이션, 비평적 사고(Critical Thinking) 어노테이션, 교육적 도움(Pedagogical Help) 어노테이션의 세 가지 유형으로 제안하였다. DLESE 어노테이션 구조는 최초의 NSDL 어노테이션 모델에서 출발하였으며, 핵심 서비스로서 명확하게 차별화된 어노테이션으로 NSDL 기술 구조와 일치한다.

DLESE 어노테이션의 메타데이터 구조는 15개 항목으로 구성되었으며, 그중에서 +표시된 7개 항목은 필수항목이다. 어노테이션 유형은 커뮤니티 이용자들의 요구를 반영한 것이며, 하나의 어노테이션에 하나의 어노테이션 메타데이터 레코드를 인정한다. DLESE 어노테이션 메타데이터의 구조는 <표 4>와 같다(Arko et al. 2006).



〈표 4〉 DLESE 어노테이션의 메타데이터 구조 필드

요소 이름	정의
+ Service	어노테이션에 대한 접근을 만드는 서비스, 단체 또는 개인명
+ Record ID	어노테이션 레코드의 식별기호
Title	어노테이션의 제목
Status	{진행중, 완성, 퇴출된 것} 중 택일
+ Date created	어노테이션 레코드 작성일
Date contributed	어노테이션 레코드 입수일
Date modified	어노테이션 레코드 최종 갱신일
+ Item ID	어노테이션의 대상인 메타데이터 레코드의 목록 번호
+ Contributor	어노테이션을 작성한 개인이나 단체
+ Type	어노테이션의 유형 {평가 권력, 편견, 다양한 청중, 설명, 편집자 요약, 교육 표준, 사례, 잘못된 생각, 양적 정보, 서평, 다른 것도 참조, 기술, 수업 비결}
+ Content	어노테이션이 부여된 대상의 내용이나 정보원에 대한 전반적인 평가
Format	{오디오, 그래픽, 텍스트 비디오} 중 선택
Context	어노테이션이 직접 부여된 자원의 페이지(URL)
More info	하나 이상의 XML 문서 형태로 제공되는 추가 정보
Share	어노테이션 작성자의 이름이나 이메일과 같은 정보를 DLESE 사용자 인터페이스에 공개할 지 여부를 표시

## 5. 어노테이션을 이용한 정보서비스 개발

### 5.1 웹 슈퍼구조(Web Superstructure)

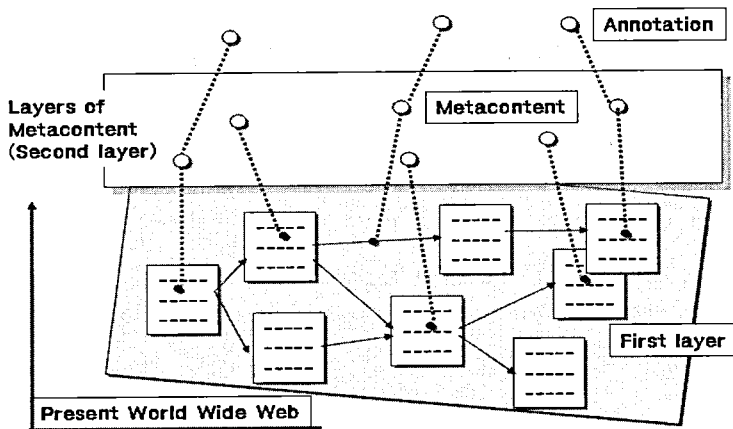
최근 인터넷에서 사용자가 직접 만들어 올린 글, 사진, 이미지, 음악, 동영상 등의 사용자 제작 콘텐츠(UCC: User Created Content)가 인기를 얻고 있다. UCC는 네티즌들이 직접 만들어 유통시킨 만큼 가벼운 일상생활의 재미나 유익한 정보를 다루는 것들이 많다. UCC의 유형은 아래와 같이 매우 다양하게 나타난다.

- 블로그의 글
- 일상을 찍은 사진과 동영상
- 인터넷 뉴스에 단 댓글
- 인터넷 쇼핑몰의 구입 후기
- 책 서평 또는 영화, 연극 등의 관람평
- 다른 네티즌의 질문에 답한 것
- 직접 그린 만화와 음악파일 등

한국인터넷진흥원(2006)의 조사에 따르면 우리나라 네티즌은 참여와 공유로 대표되는 웹2.0 패러다임의 제반 환경 측면에서 양적인 성숙단계에 도달했다고 주장하였다. 인터넷 이용자의 91.6%가 카페/커뮤니티 이용, 퍼나르기(펌) 경험, 블로그/미니홈피 운영, 댓글 작성, UCC생산 등의 참여와 공유 활동 중 하나 이상의 활동에 참여하고 있다고 분석하였다.

최근 인터넷의 핵심 트렌드인 UCC에 대한 조사에서 조사대상자의 76.2%가 UCC를 이용해 본 경험이 있으며, 특히 초기적인 이용자 생성 어노테이션에 해당하는 댓글 작성은 45.6%, UCC 생산 경험은 43.2%가 있는 것으로 나타났다. 한편 UCC의 또 다른 형태로 볼 수 있는 상품평, 이용후기 등은 인터넷상에서 입소문을 형성하여 쇼핑 구매결정에 상당한 영향을 미치는 것으로 분석되었다. 따라서 UCC가 일상생활뿐만 아니라 전문분야에서도 증가하고 있으므로 웹 서비스에서 어노테이션 작성 절차와 구조를 체계화하여 어노테이션의 재사용과 공유를 지원할 수 있는 프레임워크 구축이 요구된다.

어노테이션이 콘텐츠로부터 지식을 발견하는데 유용하다는 점에 착안하여 Nagao(2003) 연구팀은 어노테이션을 콘텐츠 자체에 내장하는 것이 아니라 원본문서에 부가적인 정보를 외부에서 어노테이션하는 웹 슈퍼구조(Web Superstructure)를 제안하였다. 전통적인 웹 구조는 2차원 평면 도식인데 이러한 2차원 도식을 복수의 평면층으로 구성된 3차원 구조로 확장하는 방법으로, 이 시스템은 콘텐츠와 메타콘텐츠(metacontent)로 구성된다. 제안된 웹 슈퍼구조는 <그림 9>와 같다.



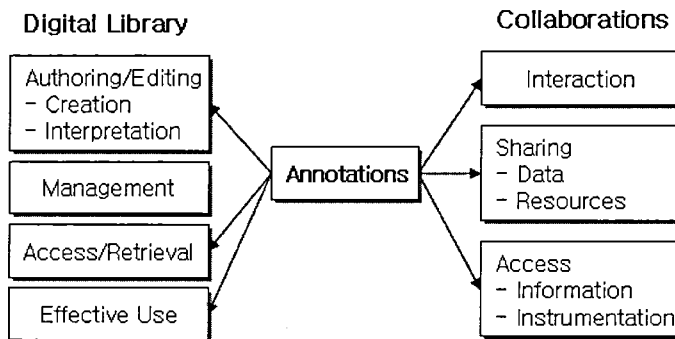
<그림 9> 웹 슈퍼구조

웹2.0 패러다임의 디지털 정보환경에서 강조되고 있는 이용자의 참여(participation)와 정보 공유를 지향하는 웹 서비스에서 어노테이션 서비스의 성공 예감은 검색포털사이트인 네이버5)의 지식검색과, 누구나 자유롭게 접근하여 편집할 수 있는 Wikipedia의 서비스에서 찾아볼 수 있다.

## 5.2 디지털도서관과 어노테이션

디지털도서관에서 어노테이션은 정보의 접근, 효과적인 이용, 데이터와 자원의 공유 등 많은 장점이 있다. 디지털도서관의 기능이 단순히 수집된 자료의 디지털화 및 원격 이용 등의 초기형태의 디지털도서관 개념에서 보다 확장되고 발전된 개념인 새로운 정보자원을 창출하고 설명해주는 개념으로 발전할 수 있도록 도와주는 요소 중의 하나가 어노테이션이라 할 수 있다.

디지털도서관 관점에서 어노테이션은 저작과 편집분야에서 새로운 자원을 만들고 주석 처리된 자료가 이용자의 이해를 돕는다. 그리고 자원에 대한 접근과 검색을 용이하게 하고 이용의 편의성을 증진한다. 협력측면에서 어노테이션은 자원과 이용자, 이용자와 이용자 사이의 상호작용을 높이고 데이터와 자원을 공유하게 하며 정보접근성을 높인다. Agosti(2004) 등이 제시한 디지털도서관에서의 어노테이션의 협력 개념은 <그림 10>과 같다.



<그림 10> 디지털도서관과 어노테이션

국내 디지털도서관에서도 외국의 DiLAS 시스템, DLESE DL 등의 어노테이션 사례를 참고하여 어노테이션에 대한 연구와 적극적인 도입이 필요하다 하겠다.

## 5.3 전자책, 전자저널과 어노테이션

정보기술과 전자출판의 발전으로 기존의 종이책이 디지털 형태로 바뀌고 있으며, 처음부터 디지털 형태로 출판되는 전자책(eBook)들도 증가 추세이다. 기존의 책이 디지털화 됨으로써 출판비용이 절감되어 저렴한 가격으로 신속하게 정보를 전달할 수 있으며, 멀티미디어 기술의 활용,

책의 저장 및 관리 편리성 등 다양한 장점을 갖는다.

또한 전통적인 책에 개인적으로 사용하던 추적인 설명, 노트, 서평 또는 밑줄과 하이라이트 표시 등의 어노테이션을 전자책에서도 활용할 수 있다. 그러나 기존의 종이책 어노테이션과 구별되는 특징은 어노테이션에 대한 접근이 용이하고 검색과 자원에 대한 공유가 가능하며 상호작용할 수 있는 점이다.

그러나 현재 대표적으로 전자책을 개발하고 서비스하는 업체인 바로북<sup>6)</sup>과 북토피아<sup>7)</sup> 등에서도 아직 초기적인 형태의 어노테이션인 게시판 형태의 독자서평과, 쓰래드기반 형식의 한줄 서평과 댓글 쓰기 지원한다. 서평에 대한 간단한 추천기능은 있지만 검색 기능 등은 찾아볼 수 없으며 서평을 작성하는 사람에게 보너스점수 등을 주어 서평 작성을 유도하고 있으나 이용이 활발하지는 못하다.

김종석(2002)은 온톨로지를 사용하는 전자책 어노테이션 시스템의 설계 및 구현 연구에서 한국 전자책 문서 표준인 EBKS(Electronic Book of Korea Standard)를 기반으로 온톨로지를 설계하여 전자책 어노테이션 구현에 사용하였다. 그러나 이 시스템은 전자책 문서 포맷의 표준화 및 저작권 문제, 어노테이션 메타데이터 표준 개발 부재 등의 이유로 적극적으로 도입되지 못하고 있다. 따라서 이러한 문제점들을 조속히 해결되어 전자책에 어노테이션 시스템이 도입되어 도서관 및 정보분야에 활용될 수 있어야 한다. 또한 과학기술분야 등 전문연구자 집단에서 집중적으로 이용하고 있는 전자저널(e-Journal)에도 어노테이션 서비스 도입이 필요하다 하겠다.

## 6. 결론 및 제언

최근 인터넷과 정보서비스의 화두는 웹2.0이다. 웹2.0의 개념은 인터넷 사용자의 참여와 정보의 공유를 특히 강조한다. 인터넷 사용자가 직접 제작하는 콘텐츠인 UCC가 인기를 끄는 것도 같은 맥락이라고 할 수 있다.

디지털 정보환경에서 사용자의 참여와 자원의 공유는 어노테이션 시스템의 활용으로 편리하고 유용하게 증진될 수 있다. 디지털 콘텐츠에 생성된 어노테이션 정보는 종이문서에서 얻을 수 있는 장점뿐만 아니라 웹 환경의 특징인 편리한 정보 접근과 이용, 자원 공유, 검색 및 재편집 등의 기능이 가능하기 때문이다. 특히 독립 어노테이션 서비스 구조에서 생성된 어노테이션 정보는 원본문서를 변형시키지 않으므로 저작권 등의 문제가 있는 디지털도서관과 전자책, 전자저널 등 여러 분야에서 다양한 형태로 적용될 수 있다.

6) <http://www.barobook>

7) <http://www.booktopia.com>

디지털환경에서 어노테이션에 관한 추가적인 연구는 대규모 디지털 콘텐츠에서 지식의 발견을 용이하게 하는 시멘틱 어노테이션 분야이다. 시멘틱 어노테이션은 의미적 정보가 주석처리된 문서를 컴퓨터가 더욱 쉽고 빠르게 이해하고 처리할 수 있도록 지원하여, 최소의 노력으로 다양한 이용자가 접근할 수 있도록 한다.

시멘틱 어노테이션 시스템은 정보 검색을 향상시키고 상호운용성(interoperability)을 증가시키는 두 가지 이점이 있으므로(Uren et al 2005) 과학기술분야와 전문적인 학술정보센터의 디지털도서관 서비스에서 우선적으로 도입할 필요가 있다. 또한 Wikipedia 서비스와 같은 전문 주제별 오픈사전(open dictionary)의 개발과, 전자책, 전자저널 등의 디지털 콘텐츠에도 어노테이션 서비스의 개발이 요구된다.

어노테이션 시스템과 결합하여 디지털 콘텐츠를 매우 다양하고 풍부하게 정보서비스를 제공할 수 있는 트랜스코딩(transcoding) 분야에 대한 연구도 필요하다. 하나의 콘텐츠를 다양한 미디어에서 이용할 수 있도록 조작(manipulation), 변형(transformation), 재사용하는 트랜스코딩은 텍스트, 오디오, 비디오 등의 멀티미디어를 재조직하여 유비쿼터스(ubiquitous) 정보환경에 효과적으로 대응할 수 있도록 도움을 주기 때문이다.

## 참 고 문 헌

- 김중석. 2002. 『온톨로지를 사용하는 전자책 어노테이션 시스템의 설계 및 구현』, 석사학위논문. 연세대학교 대학원, 컴퓨터산업시스템공학과.
- 사공철 등편. 1996. 『문헌정보학용어사전』, 서울: 한국도서관협회.
- 이재윤, 황혜경. 2006. 이용자 생성 메타데이터에 관한 연구. 『정보관리연구』, 37(3): 1-24.
- Agosti, M., et al. 2005. "DiLAS: a Digital Library Annotation Service." *International Workshop on Annotation for Collaboration*. Paris, November, 24-25.
- Agosti, M., N. Ferro, I. Frommholz, U. Thiel. 2004. "Annotations in Digital Libraries and Collaboratories-Facets, Models and Usage." *Proceedings of 8th European Conference on Research and Advanced Technology for Digital Libraries(ECDL 2004)*, Bath, UK.
- Agosti, M., N. Ferro, I. Frommholz, E. Panizzi, W. Putz, and U. Thiel. 2006. "Integration of the DiLAS Annotation Service into Digital Library Infrastructures." *Proceedings of the Workshop on Digital Libraries in the Context of Users (DL-CUBA 2006)*, 1-4.

- Annotation System for Semantic Web. [cited 2006.9.23].  
 <<http://www.ncb.ernet.in/groups/dake/annotate/index.shtml>>.
- Annotea Project Homepage. [cited 2006.9.21]. <<http://www.w3.org/2001/Annotea/>>.
- Arko, R. A., K. M. Ginger, K. A. Kastens, and J. Weatherley. 2006. "Using Annotations to add value to a Digital Library for Education." *D-Lib Magazine*, 12(5). [cited 2006.9.5].  
 <<http://www.dlib.org/dlib/may06/arko/05arko.html>>.
- CritLink : Public Web Annotation. [cited 2006.9.23]. <<http://zesty.ca/crit/>>.
- Denoue L. and L. Vignollet. :An Annotation Tool for Web Browsers and its Applications to Information Retrieval." *In Proceedings of RIAO 2000*, Paris, April 2000.
- iMarkup Annotation. [cited 2006.9.10].  
 <[http://www.imarkup.com/client/annotate\\_page.asp](http://www.imarkup.com/client/annotate_page.asp)>.
- Kastens, Kim. 2005. "The DLESE Community Review System: Gathering, Aggregating, and Disseminating User Feedback about the Effectiveness of Web-based Educational Resources." *Journal of Geoscience Education*, 53(1): 37-43.
- Kiryakov Atanas. B. Popov, L. Terziew, D. Manov, D. Ognyanoff. 2004. "Semantic Annotation, Indexing, and Retrieval." *Web Semantics: Science, Services and Agents on the WWW*, 20(10): 481-490.
- Marshal Catherine C. 1997. "Annotation: from Paper Books to the Digital Library." *Proceedings of the 2nd International Conference on Digital Libraries(DL 97)*, Philadelphia, pp.131-140.
- Nagao Katashi. 2003. *Digital Content Annotation and Transcoding*. Boston·London : Artech House.
- Nagao K., Y. Shirai, and K. Squire. 2001. "Semantic Annotation and Transcoding: Making Web Content more Accessible." *IEEE Multimedia*, 8(2): 69-8. [cited 2006.9.6].  
 <<http://citeseer.ist.psu.edu/context/1461138/320713>>.
- NSDL-The National Science Digital Library. [cited 2006.9.13].  
 <<http://www.nsd.org>>.
- Ovsiannikov Ilia A. 1999. "Annotation Technology." *International Journal of Human-Computer Studies*, 50(4): 329-362.
- Paul perry. 2001. Web Annotations. [cited 2006.8.13].  
 <<http://www.paulperry.net/notes/annotations.asp>>.

- Richard E. Rubin. 2000. *Foundations of Library and Information Science*. New York: Neal-Schuman Publishers.
- Steve B. Cousins, M. Baldonado, and A. Paepcke. *A Systems View of Annotations*. Technical Report P9910022, Xerox PARC, 2000. [cited 2006.9.11].  
<<http://citeseer.ist.psu.edu/579278.html>>.
- Uren Victoria, p. Cimiano, J. Iria, S. Handschuh, M. V. Vera, E. Motta and F. Ciravegna. 2006. "Semantic Annotation for Knowledge Management: Requirements and a Survey of the State of the Art". *Web Semantics : Science, Services and Agents on the World Wide Web*, 4(1) : 14-28.
- Wikipedia Homepage. [cited 2006.9.8].  
<[http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page)>.
- Wolfe Joanna. 2002. "Annotation Technologies: A Software and Research Review." *Computers and Composition*, 19(4) : 471-497.
- YEWAS. [cited 2006.7.25]. <<http://www.fxpal.com/people/denoue/yawas>>.
- Zellweger P., N. Bouvin, H. Jehoj, and J. Mackinlay. 2001. "Fluid Annotations in an Open World." *Proceedings of ACM Hypertext*, Denmark.