

기독교 관련 웹 사이트 내 색인에 관한 연구

A Study on Christian Website Indexing

유 영 준(Yeong-Jun Yoo)*

< 목 차 >

- | | |
|---------------------|-------------------|
| I. 서론 | IV. 웹 사이트 내 색인 작성 |
| II. 웹 사이트 내 색인 | V. 결론 |
| III. 웹 사이트 내 색인의 사례 | |

초 록

웹 사이트 내 색인은 권말색인의 기능과 유사한 역할을 한다. 웹 사이트에 접속한 후에 이용자에게 사이트 내의 특정 페이지나 보다 상세한 콘텐츠에 직접적인 접근을 제공하는 것이 웹 사이트 내 색인을 작성하는 목적이다. 웹 사이트 내 색인은 권말색인과 마찬가지로 자모순으로 배열되지만, 웹상에서는 소재위치를 표시하는 대신에 HTML의 <a> 태그에 의해서 색인 저록과 사이트의 콘텐츠가 하이퍼링크 되어 연결된다. 따라서 이 연구에서는 수작업 색인과 자동으로 표목을 생성한 후 수작업으로 통제된 색인, 두 가지 방식으로 웹 사이트 내 색인을 작성하였다. 전통적인 문헌정보학의 정보조직 수단인 권말 색인과 유사한 웹 사이트 내 색인을 도서관과 같은 사이트에서도 적극적으로 활용하는 것이 바람직하다.

주제어: 권말색인, 웹 사이트 내 색인, 수작업 색인, 자동색인, 하이퍼링크, HTML

ABSTRACT

Back-of-book-style indexes have a similar function as back-of-book indexes. The best advantage of back-of-book-style indexes for information access on the web is to give direct access to specific subjects of interest. Though back-of-book-style indexes are alphabetically arranged as back-of-book indexes, they have linked index entries to contents on the site by using an anchor tag of HTML. In this research, I have created back-of-book-style indexes in two separated ways, by hand-crafted and semi-automatic indexing. We have utilized back-of-book-style indexes, that is similar to back-of-book index of traditional information organization method of library and information science, in library circumstances.

Key Words: Back-of-Book Indexes, Back-of-Book-Style Indexes, Hand-Crafted Indexing, Automatic Indexing, Hyperlink, HTML

* 연세대학교 문헌정보학과 시간강사(yo6370@hanmail.net)

• 접수일: 2007년 10월 30일 • 최종심사일: 2007년 11월 26일 • 최종심사일: 2007년 12월 24일

I. 서론

일반적으로 도서관에서 이용자들이 자신이 원하는 문헌을 검색해서 찾아 낸 이후에 그 문헌의 특정한 내용을 다시 찾아야 하는 경우가 발생할 수 있다. 단행본의 경우처럼 분량이 많은 문헌들은 목차나 권말색인을 활용하는 것이 유용한데, 특히 검색된 문헌에서 매우 구체적이고 상세한 내용에 접근하기 위해서는 권말색인을 활용하는 것이 가장 간편한 방법이라고 할 수 있다.

권말색인은 분량이 많은 문헌에서 원하는 정보에 직접적으로 접근하기 위한 접근점으로서의 역할을 수행하는 주제나 이름이 체계적으로 배열된 저록들을 지칭하는 용어다. 권말색인은 전통적으로 자모순으로 배열되지만, 시소러스에서 제공되는 계층적 배열도 사용할 수 있고, 분류체계의 경우에서처럼 전체가 계층적일 수도 있다. 그리고 권말색인의 유무가 그 문헌의 가치를 평가하는 요인으로도 작용할 수 있다. 이러한 전통적인 도서의 경우에서처럼 웹상에서도 권말색인과 같은 검색 도구를 이용 할 수 있어야 한다.

현재 웹은 많은 사람들에게 주요한 정보원이 되고 있다. 웹 이용자들은 자신들이 관심을 가지고 있는 특정 주제와 관련된 웹사이트를 쉽게 찾을 수도 있고, 그렇지 않을 수도 있다. 그러나 이용자들은 웹사이트에 접속을 하고 난 후에 자신들이 필요로 하는 특정 정보를 찾기 위해서 오히려 웹사이트를 검색할 때 보다 더 많은 노력과 시간을 드리는 경우도 있다. 이러한 문제를 해결하기 위한 방법으로는 사이트 맵이나 사이트 내 검색을 활용하는 것이 우리 웹 환경에서는 일반적이다.

그러나 웹 사이트 내 색인은 디렉토리 서비스의 단순한 형태인 사이트 맵이나 사이트 내 검색과 차별화되는 점이 있다. 사이트 맵에서는 문헌의 목차와 같은 기능은 제공할 수 있지만 보다 상세한 정보를 검색하는 데에는 한계가 있어서 이용자가 원하는 상세하고 구체적인 정보를 찾으려면 다시 시간과 노력을 들여야 한다. 사이트 내 검색의 경우에서도 이용자들이 웹에서 경험하는 것과 유사하게 이용자가 원하는 정보를 정확하게 검색해 내는데 한계가 있다. 다시 말하면 개념을 식별하지 못하므로 관련 없는 정보들이 검색 결과에 포함되어 적합성이 떨어지는 경우가 빈번하다.

더불어 구글(Google)에서 해당 사이트에 존재하는 자원만을 대상으로 검색하는 기능이 있으므로 별도로 검색 엔진을 만들지 않아도 된다. 예를 들어 W3C(WWW Consortium) 사이트 내에서 'html'에 관한 정보를 검색하려면, 구글의 검색 창에 'site:http://www.w3.org html'와 같이 입력하면 W3C 사이트 내의 'html'과 관련된 자원의 결과만을 얻을 수 있다.

따라서 웹 사이트 내에서 이용자가 원하는 매우 특정한 정보에 직접 접근하고 보다 정확하고 빠르고 간편하게 검색할 수 있는 방법을 고려해야 한다. 이러한 문제들을 해결 할 수 있는 방법이 바로 이미 문헌정보학에서 활용하고 있는 수단인 권말색인이다.

단행본의 권말색인은 개인 웹사이트, 이미지나 영상, 사운드와 같은 비텍스트 자료(non-textual material) 그리고 텍스트와 다른 매체가 결합된 멀티미디어 자원뿐만 아니라 웹 사이트 내의 문서

들의 탐색에도 활용될 수 있다. 하지만 웹 사이트에서 권말색인의 개념은 단행본에서의 권말색인과는 차이가 있는데, 색인의 대상이 HTML 파일이며 이 파일에 링크가 포함될 수도 있다. 따라서 개개의 색인 저록이나 표목이 사이트 내의 웹 페이지나 페이지의 상세한 내용으로 직접 하이퍼링크(hyperlinked)되는 자모순 권말 색인(alphabetical back-of-the-book style index)을 웹 사이트 내 색인이라고 정의할 수 있다. 그래서 이 논문에서는 기존의 권말색인과 웹 사이트의 색인을 간단히 비교 하였고, 이미 개발되어 운영되고 있는 사례를 검토하였으며, 웹 사이트를 대상으로 색인 시스템을 구축하였다.

이 연구의 목적은 웹 사이트의 콘텐츠를 검색하는 수단인 사이트 내 검색과 사이트 맵의 단점을 보완할 수 있는 웹 사이트 내 색인을 구축하는 것이다. 즉 사이트 내 검색이 가지고 있는 문제인 낮은 적합성과 사이트의 콘텐츠에 대한 직접적인 접근을 제공하는데 한계가 있는 사이트 맵의 단점을 보완하여, 웹 사이트의 콘텐츠에 대한 직접적인 접근법과 색인어 브라우징을 제공하는 웹 사이트 내 색인시스템을 제안하는 것이 이 연구의 목적이다.

그리고 웹 사이트 내 색인을 작성하기 위해서 참조할 만한 색인 작성 지침을 조사한 결과, ISO 5963¹⁾과 ISO 999²⁾에서는 일반적인 색인에 관한 표준이고 직접적으로 웹 사이트 내 색인에 관한 세부 사항은 명시되어 있지 않았다 그리고 국내 표준으로는 1999년에 한국정보통신기술협회(TTA)에서 제정한 정보통신 단체표준인 '색인작성지침'(TTA.KO-11.0012)을 검토하였으나 외국의 여러 표준을 종합하여 정리한 정도의 지침이었다. 또한 김효열³⁾은 도서 권말색인의 작성을 위한 표준의 필요성을 제시하면서, 권말색인에서 표목을 자동으로 생성할 수 있는 시스템이 개발된다면 표목의 망라성을 확보할 수 있다고 주장하였다. 그리고 자동으로 생성된 표목에서 저자가 수작업으로 표목에 통제를 가한 후에 색인 작성 지침에 따라 편집하면 더 나은 색인이 될 수 있다고 주장하였다.

따라서 이 연구에서의 연구 방법으로는, 위의 선행연구에서도 밝혀진 것과 같이 수작업으로 권말 색인을 작성하는 것이 가장 주요한 방법이고, 자동으로 표목을 생성하는 것이 보완적인 연구 방법으로 활용되었다. 이 연구에서는 연구 대상인 '성약출판사' 사이트의 각각의 페이지를 모두 접속하여 분석하였고, 수작업으로 적절한 색인어를 선정한 후, 2차 선정과정을 거쳐서 최종적으로 219개의 색인어를 대상으로 '성약출판사'의 웹 사이트에 대한 색인을 작성하였다.

두 번째 분석대상은 종교개혁이후에 처음으로 작성된 『하이델베르크 요리문답』 1563년 판의 번역본⁴⁾인데, 이 문헌의 본문을 형태소 분석기를 사용하여 자동으로 색인 표목을 생성하고 이 색인어들 중에서 수작업으로 의미 있는 색인어 121개를 다시 선정하여 『하이델베르크 요리문답』 웹

1) ISO 5963. *Documentation -- Methods for examining documents, determining their subjects, and selecting indexing terms*(1985).

2) ISO 999: 1975. *Documentation: Index of a Publication*.

3) 김효열, 도서권말색인의 작성지침과 자동생성에 관한 연구(석사학위논문, 연세대학교 대학원 문헌정보학과, 1995).

4) 독립개신교회 교육위원회 옮김, 하이델베르크 요리문답(서울 : 성약출판사, 2004).

사이트 내 색인을 작성하였다. 또한 웹 사이트 내 색인을 구축하기 위해서 요리문답의 본문 전체를 HTML 파일로 변환하였고, 온라인 도서의 형태로 홈페이지를 구축하였다.

II. 웹 사이트 내 색인

1. 권말 색인과의 비교

웹 사이트 내 색인은 자모순 권말색인과 유사한 웹 색인의 일종으로, 색인 저록들이 적절한 웹 페이지나 페이지의 일부분에 직접적으로 하이퍼링크 된다. 이러한 웹 사이트 내 색인이 전통적인 권말색인과 가장 두드러지게 다른 점은 웹상에서 색인의 기능을 수행해야 하기 때문에 웹 환경에 적합하게 색인과 웹 페이지의 본문이 링크되어야 한다는 것이다. 그리고 개개의 웹 페이지나 검색 도구, 내부 데이터베이스에 의한 검색을 위해서 통제키워드에 기반한 색인이어야 한다. 이러한 웹 사이트 내 색인은 특정한 사이트 전체, 하위 사이트, 인트라넷, 온라인 논문들의 집합, 온라인 교과 과정이나 전자도서의 색인이 될 수도 있다. 따라서 이러한 웹 사이트 내 색인을 작성한다는 것은 개개의 웹 사이트, 인트라넷, HTML 문서들의 집합이나 웹 사이트들의 집합을 위한 색인을 생성하는 것을 의미한다.

기존의 권말색인에서 중요한 점은 웹 사이트 내 색인 작성에서도 동일하게 중요하다. 어떤 개념들이 색인할 가치가 있는가를 결정하는 분석적인 과정인 색인 작성(indexing)은 어느 저록들(entries)을 이용할 것인가와 그 저록들을 어떻게 배열할 것인가와 관련되어 있다. 따라서 웹 사이트 내 색인은 색인 작성 기술에 능숙한 사람들에 의해 가장 잘 작성될 수 있다. 비록 개별적으로 수작업에 의해 작성된 메타데이터로부터 색인을 자동으로 생성하는 경향이 증가하고는 있지만, 웹 사이트 내 색인은 컴퓨터에 의해서가 아니라 대부분 사람들에 의해서 만들어져 왔다.⁵⁾ 전문적인 색인자, 사서와 웹 사이트 설계자들이 작성한 색인에서는 개념을 식별하기 위해서 텍스트를 분석하고, 훈련을 받은 색인자들은 세목들과 상호 참조를 생성하기 위해서 주제 간의 관련성을 파악할 수도 있다. 그리고 색인자들은 이용자들이 정보를 어떻게 찾을지를 예상해서 다수의 동의어들도 제공할 수 있다.

웹 사이트 내 색인을 자동으로 생성하는 소프트웨어를 사용할 수도 있지만, 그 결과가 색인자가 수작업으로 작성한 색인보다 그 결과가 만족스럽지 못한 것이 사실이다. 그리고 자동 색인 소프트웨어만으로는 개념들을 식별할 수 없으며, 이들은 텍스트 내에서 구분되지 않는(undifferentiated) 단어들을 색인할 수 있다. 따라서 전문적인 교육을 받은 색인자가 웹 사이트 내 색인을 작성하는

5) Browne, Glenda and Jonathan Jeremy. *Website Indexing: enhancing access to information within websites*. 2nd ed.(Adelaide : Auslib Press, 2004),

것이 가장 좋은 방법이라는 것은 기존의 권말색인의 경우와 동일하다.⁶⁾ 그러나 개념을 식별할 수 없는 자동 색인 소프트웨어의 단점을 개념을 구분할 수 있는 색인 전문가로 하여금 보완하도록 하여, 자동으로 생성된 표목에서 적절하게 표목을 선정하도록 하는 것도 유력한 방법이 될 수 있다.

그리고 권말 색인과 웹 사이트 내 색인은 공통점과 차이점을 가지고 있다. 일반적으로 색인을 작성할 때 필요한 과정 즉 색인 대상 문헌이나 사이트의 주제적 특성을 분석하여 개념을 추출하고, 그 개념을 용어로 표현하는 과정과 자모순으로 색인 저록을 배열하는 기본적인 구조 등 대부분의 체계는 동일하다. 그러나 차이점은 주로 색인 대상의 차이에서 생기는 것이 대부분이다. 첫째, 색인 대상이 도서의 경우에는 크기는 페이지나 단어의 수를 계산하지만, 웹 사이트의 경우에는 웹 페이지들의 수나 단어의 수 외에도 파일의 메가바이트 수를 고려해야 한다. 둘째, 색인이 제공되는 형식(format)이 웹 사이트 내 색인에서는 더 중요하다. 그 이유는 웹 사이트 내 색인에서는 HTML 색인이기 때문에 웹 페이지들 안에 갱신을 위한 메타데이터가 포함된 형식이어야 하는지를 결정해야 한다. 셋째, 두 번째와 연계해서 고려해야 할 부분으로 도서의 권말 색인은 새로운 판이 나오지 않는 한 갱신되지 않지만, 웹 사이트 내 색인은 주기적으로 갱신되어야 한다. 넷째, 권말 색인과는 달리 웹 사이트 내 색인의 경우에는 자모순 색인 화면의 맨 위에서 자모순으로 배열된 특정 자모의 색인 저록으로 직접 이동할 수 있다. 다시 말해서 화면 내에서 <a> 태그를 이용하면 스크롤바를 이용하지 않고도 화면에서 보이지 않는 아래쪽의 특정 글자, 예를 들어 “하” 영역으로 직접 이동하는 것이 가능하다. 다섯째, 권말색인에서는 본문의 페이지가 명시되어야 하지만, 웹 사이트 내 색인에서는 링크로 직접 연결되기 때문에 소재 정보를 표시할 필요가 없다.

위에서 비교한 내용을 표로 제시하면 아래의 <표 1>과 같다.

<표 1> 권말색인과 웹 사이트 내 색인의 차이점

내용 \ 종류	권말색인	웹 사이트 내 색인
본문/컨텐츠와 연결 방법	소재정보	하이퍼 링크
크기 표시 기준	전체 페이지 수	웹 페이지 수나 단어의 수 또는 파일의 메가바이트 수
색인 제공 형식	중요성이 낮음	메타데이터 포함 유무를 결정해야 하므로 중요성이 높음
갱신 주기	새로운 판이 나올 때	주기적으로 계속
특정 색인 저록으로의 이동	처음부터 검색	직접 이동

2. 웹 사이트 내 색인의 장점

위에서 제시한 차이점 및 공통점과 더불어 웹 사이트 내 색인이 가지는 장점을 제시하면 다음과

6) Ibid.

같다. 첫째, 관심 있는 특정 주제들에 대한 직접적인 접근을 가능하게 하여 적합한 특정 페이지나 상세한 내용을 찾을 수 있고 시간을 절약할 수 있다. 둘째, 특정 사이트의 전체 콘텐츠의 개요를 브라우징 할 수도 있다. 따라서 이미 알고 있는 용어나 내용을 찾는 것만이 아니라 그 사이트에서 어떤 정보들을 다루고 있는지를 확인할 수 있다. 그리고 그 주제를 기술하는데 사용된 용어가 어떤 것인가와 상관없이 특정 주제에 관한 모든 용어들에 대한 링크가 범주화되어 있어서 특정한 주제를 기술한 용어들이 아무리 많다고 하더라도 그 용어들은 세목들(subheadings)에 의해 구분된다. 그리고 이 세목들은 이용자들의 질문에 대해서 해답을 줄 수 있는 가능성이 높은 단서들을 제공한다.

웹 사이트 내 색인은 특정 주제의 주요한 논점들을 제시하고, 사이트 내 검색에서 정교하지 않은 탐색어의 사용 즉 이용자가 즉흥적으로 선택하는 탐색어들로 인해 발생할 수 있는 어려움들을 피할 수 있게 해준다. 또한 상호참조는 정보를 찾을 수 있는 웹 사이트 내의 다른 색인 저록이나 표목도 알려준다.

그리고 웹 사이트 내의 정보에 접근하는데 있어서 우리에게 익숙한 형식을 제공한다는 것이 웹 사이트 내 색인의 장점 중의 하나가 될 수 있다. 웹 사이트 내 색인은 사이트의 크기가 어느 정도 크거나 복잡할 때 특히 활용가치가 높지만, 동시에 색인을 작성하는 것이 그 만큼 더 어려워지는 것도 사실이다. 해당 사이트에서 구조화된 링크들에 의해서 이용할 수 있는 디렉토리 주제 탐색과 사이트 내의 탐색 엔진에 의한 키워드 탐색을 보완하는 것이 웹 사이트 색인의 목적이라고 할 수 있다. 특히 사이트 맵과 비교할 때, 목차보다 더 많은 표목들을 제시할 수 있는 색인들이 보다 상세한 정보 단위에 대한 접근을 제공할 수 있다.

따라서 색인은 웹 사이트의 유용성을 제고할 수 있는 중요한 정보탐색 도구로서 간편한 탐색을 제공한다. 이용자 자신들의 어휘를 이용해서 내용에 대한 접근점을 제공하고 본문에서 다루어진 개념들에 대한 접근도 제공한다. 계층구조에서 많은 수의 링크들을 따라다니면서 찾아다닐 필요 없이, 정보의 상세한 단위(granular chunks of information)들에 대한 직접적인 접근을 제공하는 것이 웹 사이트 내 색인의 가장 중요한 점이다(Liese 2004).⁷⁾

이러한 웹 사이트 내 색인 자체의 장점과 더불어 웹상에서 정보에의 접근성과 관련해서 차모순으로 배열되는 웹 사이트 내 색인의 장점은 다음과 같다(Web Indexing SIG Brochure - United States).

- 모든 이용자에게 친숙한 이용자 친화적 색인
- 동의어들과 다양한 유형의 철자를 제시함
- 표목과 <a> 태그로 구체적인 내용을 더 정확하게 색인함
- 사이트 내의 링크로 인해 검색 엔진 사용의 극대화에 기여함
- 계층적인 디렉토리를 사용하는 것과 비교해서 한 번의 클릭으로 내용에 접근

7) Leise, Fred, 2004. "Improving usability with a website index." [online]. [cited 2007.2.8]. <<http://www.web-indexing.org/article-leise.htm>>.

- 사이트의 유용성을 전반적으로 개선함

이용자가 보다 간편하게 손쉽게 자신이 원하는 다양한 용어들을 훑어본 후에 해당 용어를 클릭 하면, 특정 페이지나 페이지의 상세한 내용에 직접적으로 접근함으로써 그 웹 사이트의 이용성을 제고시키며 이용자가 다시 사이트를 방문하도록 유도할 수 있다.

그리고 이와 같은 장점이 있는 자모순 색인을 적용하기에 가장 적합한 사이트의 특징들을 몇 가지 제시하면 다음과 같다(Web Indexing SIG Brochure - United States).

- 중간 크기(25-500 pages)의 웹 사이트
- 대형 사이트의 콘텐츠가 자주 바뀌지 않는 정적인 섹션
- 콘텐츠가 풍부한 사이트
- 반복적 방문자들이 접속하는 사이트; 회원제로 운영되는 단체들이나 교육기관, 도서관, 금융 서비스를 제공하는 기관 등
- 인트라넷의 하위 사이트(Intranet subsites)
- 온라인 연속간행물이나 전자저널(Online periodicals or e-zines)
- E-learning/ 온라인 교과과정(online courses)

따라서 지나치게 크지 않으면서 콘텐츠가 빈번하게 변하지 않고, 단편적인 콘텐츠 보다는 내용에 대한 설명이 풍부한 사이트들이 자모순 웹 사이트 내 색인을 적용하기에 적합하다. 그리고 이용자들이 반복적으로 접속하여 이용하는 사이트인 교육기관과 도서관 등에 적합한데, 이와 같은 자모순 웹 사이트 내 색인을 구축하여 제공하고 있는 도서관이나 교육기관의 사이트를 찾기가 거의 불가능한 것이 우리의 현실이다. 이러한 현상이 나타나는 것은 문헌정보학의 유용한 정보 조직 수단을 웹상에서 활용하려는 새로운 연구가 활발하지 못하고, 색인이나 권말색인에 대한 관심이나 연구가 부족한 것이 가장 본질적인 이유라고 할 수 있다. 먼저 문헌정보학 분야에서 출판되는 단행본의 권말 색인 유무나 작성 실태에 대한 분석적인 연구가 선행되어야 하고, 분류와 목록의 경우처럼 교육과정에서도 보다 실제적인 색인 작성 교육이 필요하다.

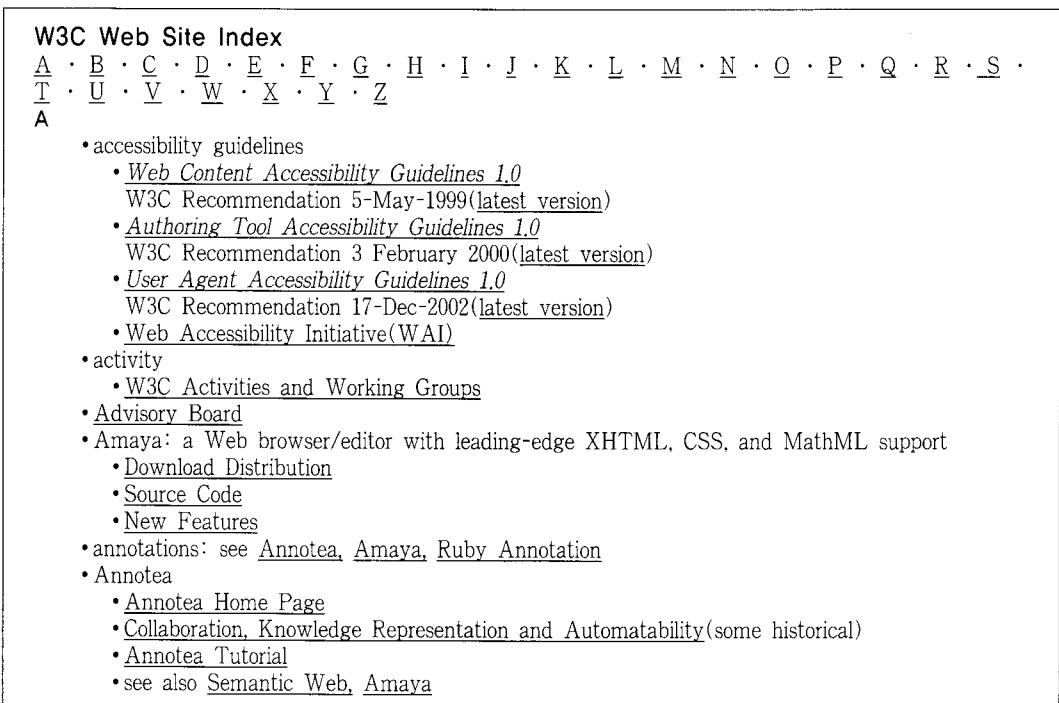
Ⅲ. 웹 사이트 내 색인의 사례

1. W3C(WWW Consortium) 웹 사이트 내 색인(www.w3.org/Consortium/siteindex)

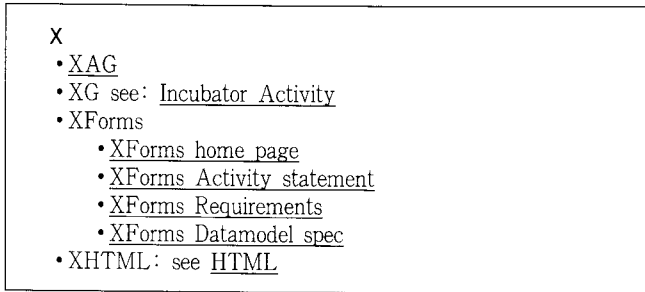
W3C 사이트에서는 'W3C Search, Site Index, Keywords'라는 제목으로 사이트 내의 정보를

탐색하는 세 가지 방법을 한 페이지에 모아서 서비스를 제공하고 있다. 즉 키워드를 입력하여 탐색하는 방법, 웹 사이트 내 색인과 용어집과 유사한 방식의 서비스가 제공되고 있다. 이 중에서 사이트 색인의 일부분을 <그림 1>에 제시하였고, 설명은 다음과 같다. 이 사이트의 색인은 특정 알파벳으로 시작하는 색인으로 직접 이동할 수 있도록 페이지 내에서 <a> 태그를 적용하고 있다. 예를 들어, 'X'를 클릭하면, <그림 2>와 같이 'X'로 시작하는 색인 저록들로 이동하게 되는 것이다. 이와 같이 페이지 내에서 <a> 태그를 사용하는 이유는 색인의 크기가 커서 한 번에 모든 색인을 브라우징 할 수 없기 때문이다. 그리고 색인의 구조는 세목이 없는 주표목에는 링크를 걸어놓고 있지만, 계층구조로 색인된 레코드의 경우에는 실제로 사이트 내의 특정 페이지나 상세한 내용으로 연결되는 하이퍼링크는 세목에만 적용하며 주표목은 세목들을 범주화하기 위한 기능만 할 뿐이다. 예를 들면, 'accessibility guidelines' 나 'Annotea' 와 같은 주표목은 하이퍼링크가 없고, 세목에만 하이퍼링크를 적용하고 있다.

사이트 내의 특정 페이지나 특정 페이지의 상세한 내용으로의 직접적인 접근을 모두 제공하는 것을 확인할 수 있었다. 또한 '보라'(see) 참조와 '도보라'(see also) 참조를 활용하여 이용자에게 다양한 접근점을 제공함으로써 탐색의 정확성을 향상시키려는 노력을 하고 있었으며, 보통명사와 고유명사를 구분하지 않고 자모순에 의한 단일 기준으로 통합하여 배열하고 있었다.



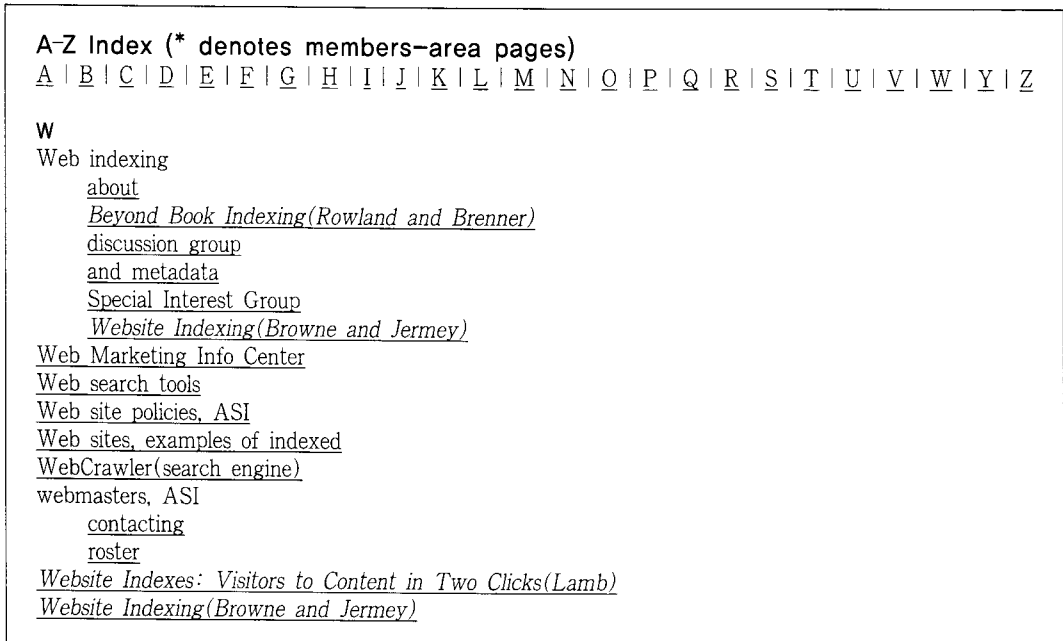
<그림 1> W3C 웹 사이트 내 색인의 예



〈그림 2〉 'X'로 시작하는 색인 저록으로 이동한 결과

2. 미국색인자협회(American Society of Indexers) 웹 사이트 색인
(www.asindexing.org/site/backndx.htm)

아래에 제시한 〈그림 3〉은 미국색인자협회의 사이트 색인으로 W3C 사이트와 다르게 탐색 기능이나 용어집을 활용해서 사이트의 특정한 내용을 찾을 수 있는 서비스는 제공되지 않는다. 그러나 W3C 사이트 색인처럼 특정 알파벳으로 시작하는 색인 저록에 직접 접근할 수 있도록 색인 페이지 내에서 <a> 태그를 사용하였으며, 색인저록들이 계층구조로 표현되었고, 인명을 색인어와 통합 배열한 것도 동일하다.



〈그림 3〉 미국색인자협회의 사이트 색인의 예

그러나 W3C 사이트 색인과 구별되는 특징이 있다. 먼저 ‘도보라’ 참조는 전혀 사용하지 않았으며, ‘보라’ 참조는 두문자어(acronym)와 완전명(full name)을 연결하는 데에만 사용하였고, 자원의 제목과 그 저자의 성을 포함시켜서 해당 자원에 연결되도록 하고 있다. 이러한 경우에도 자원의 유형이 단행본인지 학술지에 게재된 논문인지 등을 식별할 수 있도록 구분하여 표시하였다. 이와 연관된 내용을 예를 들어 설명하면, “*Website Indexing(Browne and Jeremy)*”와 같은 표목은 웹 자원의 제목과 저자를 나타내는 것으로 글자를 기울임꼴로 표시하여 다른 색인 기록들과 구분하고 있다. 그리고 예로 든 자원은 인쇄형태로 출판된 도서가 아니라 웹에서 PDF 파일 형식으로 출판된 웹 자원인데 단행본 형식이라는 것을 표시하고 있다. 이 색인 기록을 클릭하면, 아래와 같은 간단한 서지정보를 보여주며 파일형식으로 된 이 자원을 구입할 수 있는 웹 사이트로 연결된다.

Browne, Glenda and Jonathan Jermeý. *Website Indexing: Enhancing Access to Information Within Websites*. 2nd ed. Adelaide, South Australia: Auslib Press, 2004. ISBN 1875145567. [Purchase this book online](#) or as a [PDF](#).

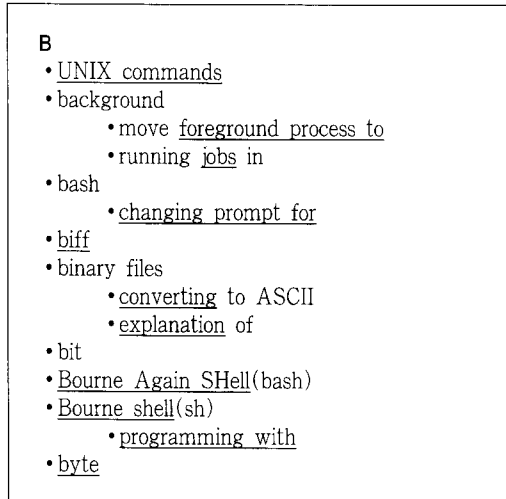
또한 동일한 제목의 자원이 다른 색인 기록에서도 출현하는 것을 확인할 수 있다. 즉 “*Website Indexing(Browne and Jeremy)*”이 ‘Web indexing’의 제목으로 기술된 것을 알 수 있다. 이러한 것은 동일한 색인 표목을 연관되는 여러 색인 기록에 동시에 기술함으로써 이용자로 하여금 여러 통로를 통해 그 자원에 접근할 수 있는 방법을 제시한 것이라고 할 수 있다. 그리고 이 색인은 웹상에서 하이퍼링크된 표목에 밑줄이 생기지 않도록 표시하였으나, 이 논문에서는 구분을 위해서 하이퍼링크된 표목에 밑줄을 표시하여 구분하였다.

3. UNIX 매뉴얼(manual) 웹 사이트 내 색인(unixhelp.ed.ac.uk/index/index.html)

〈그림 4〉는 UNIX 매뉴얼을 웹 사이트에서 온라인으로 이용할 수 있도록 작성된 것으로, 한 권의 온라인 도서인 UNIX 매뉴얼에 대한 색인을 작성한 경우이다. 따라서 앞에서 제시한 색인 시스템과는 성격이 다른 웹 사이트 내 색인이라고 할 수 있다. 다른 웹 사이트 내 색인과 다르게 이 색인은 색인 화면에 알파벳만이 제시되어 있고, 특정 알파벳을 누르면 그 알파벳으로 시작하는 색인 기록들이 새 화면에 나타난다. 특정 알파벳을 클릭하면 주로 UNIX 시스템을 사용하는 관리자들이 필요로 하는 명령어들과 관련 용어들로 연결되어, 그 명령어의 정의와 사용법에 대한 설명이 화면에 나타난다. 그리고 ‘보라’ 참조와 ‘도보라’ 참조를 사용하고 있다.

이 색인은 도서 한 권을 웹 사이트로 설계하고 그 도서의 내용에 대한 색인을 작성함으로써 인쇄 형태의 매뉴얼을 이용할 때 보다 훨씬 용이하게 매뉴얼을 활용할 수 있다는 장점이 있다. 당연히

도서의 크기 즉 내용의 분량이 대단히 큰 경우에 적용할 수 있는 방법이라고 할 수 있다.



〈그림 4〉 UNIX 매뉴얼 색인의 예

이 외에도 미국의회도서관과 같은 대형 도서관이나 시카고 공공도서관, 그리고 캠브리지 대학과 같은 교육기관, 미국의 정부기관들에서 자신들의 사이트의 유용성을 높이기 위해서 웹 사이트 내 색인을 제공하고 있었다. 그러나 앞에서 자모순 웹 사이트 내 색인을 적용하기에 적당한 사이트의 크기 즉 웹 페이지의 수를 제시한 것처럼 지나치게 큰 사이트 색인의 경우에는 그 사이트 내의 특정 페이지로의 접근이 대부분이고, 특정 페이지의 상세한 내용에 접근하는 사례는 많지 않았다. 하지만 크기가 큰 웹 사이트의 경우에는 사이트 맵으로 그 사이트의 특정 페이지로의 연결을 제공할 수 없고 사이트 내 검색의 효율성이 크기가 작은 사이트 보다 더 떨어지기 때문에 해당 사이트의 검색의 용이성과 정확성 그리고 사이트 활용의 유용성 제고를 사이트 색인을 통해서 해결하고 있음을 알 수 있었다.

IV. 웹 사이트 내 색인 작성

1. 성약출판사 웹 사이트(www.sybook.org) 색인

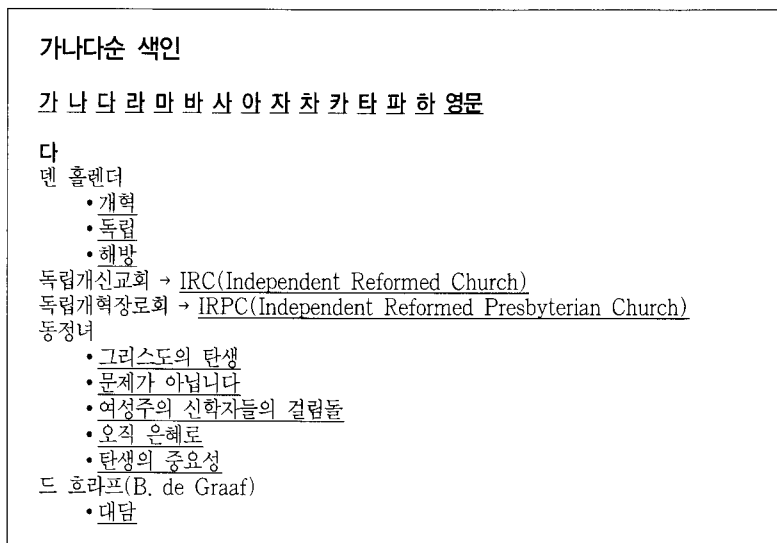
이 사이트를 색인 대상으로 선정한 이유는 연구자가 해당 출판사에서 발행되는 도서에 대한 풍부한 지식을 가지고 있기 때문이다. 즉 이 웹 사이트의 컨텐츠에 대해서 매우 잘 알고 있기 때문에

색인어를 선택함에 있어서 보다 정확성을 기할 수 있다고 판단되었다.

‘W3C 웹 사이트’나 ‘미국색인자협회 웹 사이트’와 유사한 ‘성약출판사’ 사이트에서 수작업으로 색인 표목을 선정하였다. 앞의 2.2에서 제시한 내용 즉 ‘자모순 색인에 적합한 사이트’에 근거했을 때, 이 사이트는 대략 155페이지 정도로 중간 크기의 웹 사이트에 해당하며, 이 사이트에서는 단행본의 목차와 간단한 소개, 그리고 개혁신학과 연관된 간략한 글이나 논문 등 다양한 내용이 실려 있어서 풍부한 콘텐츠를 이용할 수 있다. 그리고 이 출판사에서 발행하는 서적의 이용자는 일반적인 기독교 신자보다는 신학자나 목회자 또는 보다 전문적으로 깊이 있게 성경을 연구하려는 이용자들이 주로 반복적으로 접속하는 사이트이다. 또한 새로운 도서가 출판되거나 새로운 글이나 논문이 게재될 때까지 걸리는 시간적인 간격이 긴 편이어서 페이지를 갱신하는 간격도 길다. 이러한 근거에 의해서 웹 사이트 내 색인을 작성하는데 성약출판사 사이트가 적합하다고 연구자는 판단하였다.

선정한 색인어는 219개 이었고, 178개의 색인어는 하이퍼링크로 연결되었으며 나머지 41개의 색인어는 세목을 전개하기 위한 주표목이거나 동의어를 연결하기 위해서 사용된 ‘보라’ 참조의 목적으로 사용된 색인어들이었다. 그리고 색인의 일부를 <그림 5>에 제시하였다.

이 웹 사이트의 색인은 한 화면으로만 구성하였다. <그림 5>에서 제시된 것처럼 화면의 상단부에 ‘가’부터 ‘영문’까지의 순서로 배열된 글자 각각에 페이지 내의 <a> 태그를 적용하여 해당 글자를 클릭하면 그 글자의 자음으로 시작하는 색인 부분으로 연결된다. 그리고 해당 색인 부분에서 모음순으로 배열된 색인 저록에서 원하는 색인 저록이나 표목을 클릭하면 성약출판사 사이트 내의 특정 페이지나 특정 내용으로 안내된다.

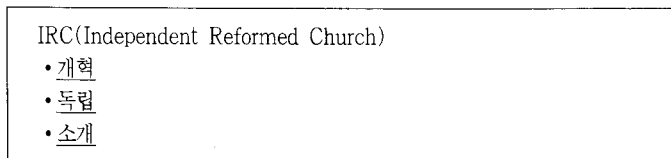


<그림 5> 성약출판사 웹 사이트 내 색인의 일부

이 웹 사이트에는 도서에 관한 정보와 개혁신앙과 관련된 내용의 글들이 주를 이룬다. <그림 5>의 색인 저록에 제시된 색인 저록들에 대해서 설명하면 다음과 같다. '덴 홀렌더'는 주표목으로 '개혁(55)', '독립(55)', '해방(55)'의 세목을 전개하기 위한 것이며, 괄호 안에 있는 '55'라는 숫자는 소식지의 번호를 의미한다. 이 사이트에서는 소식지에 강해설교나 논단 또는 교회소식 등이 실린다.

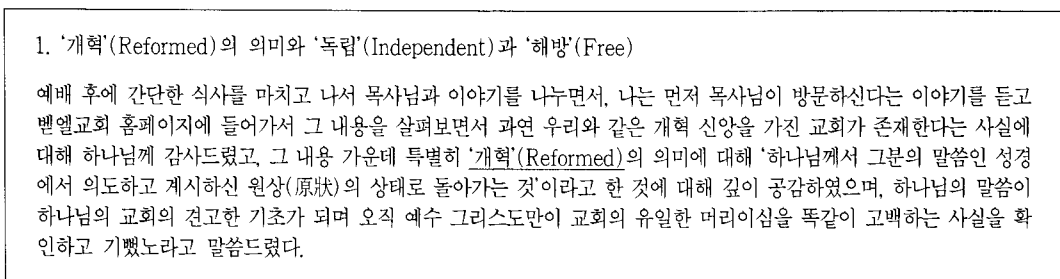
아래에 제시되는 구축 사례들은 단일 주제어는 항상 단일 페이지로의 접근된다는 점을 상기시키는 것이 필요하다고 생각한다.

'보라' 참조는 '→'를 사용해서 표시하였다. 즉 '독립개신교회'에 관한 내용을 보려고 하는 이용자는 'IRC(Independent Reformed Church)'를 클릭하면 해당 색인 저록으로 연결된다. <그림 6>은 Independent Reformed Church)'를 클릭했을 때 나타나는 색인 저록이다. 그리고 '보라' 참조에서 '→' 기호를 사용한 것은 한국목록규칙 4판의 부록에 있는 '용어 해설'의 방식을 따른 것이다.



<그림 6> 'IRC(Independent Reformed Church)' 색인 저록

예를 들어 설명하면, 위의 <그림 6>의 세목 중에서 '개혁'을 선택하면, 아래와 같이 '개혁'이 포함된 글로 이동하게 된다.



그래서 본문 중에서 '개혁'이라는 용어의 정의를 이용자가 파악할 수 있게 되는 것이다.

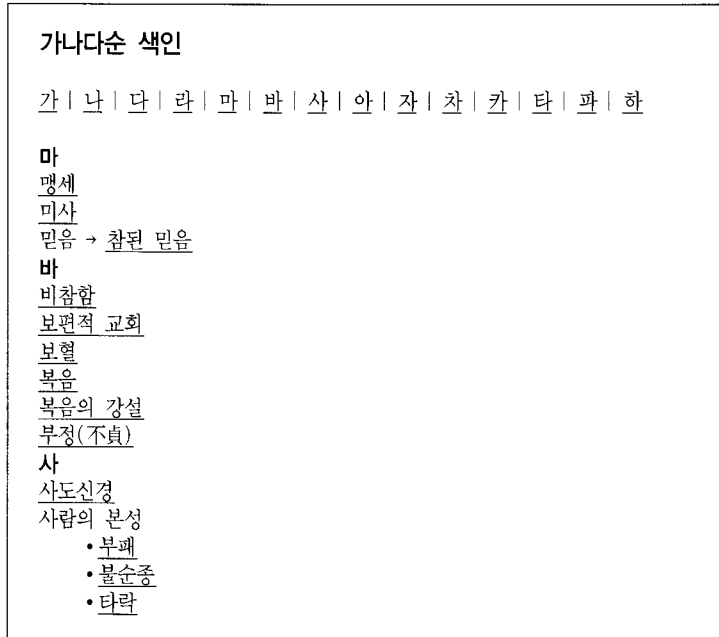
2. 하이델베르크 요리문답 웹 사이트 내 색인

이 사이트는 '3.3 UNIX 매뉴얼 웹 사이트 내 색인'과 같은 온라인 도서에 대한 색인으로, 색인어

는 자동으로 생성하였으며, 웹 사이트 내 색인을 구축하려는 목적으로 연구자가 만든 웹 사이트이다. 색인어 추출의 대상이 된 하이델베르크 요리문답은 종교개혁의 영향을 받아 작성된 신앙고백서인데, 여기에서는 성약출판사에서 발행한 번역본을 활용하였다.

이 번역본의 본문을 텍스트 파일로 작성한 후에 형태소 분석기를 이용해서 색인어를 생성하였다. 생성된 색인어 중에서 의미 있는 색인어를 다시 선정하여 121개의 색인어를 얻을 수 있었다. 이 색인어들을 웹 사이트 내 색인으로 작성하였고, 이 중에서 103개는 하이퍼링크로 연결되었고, 나머지 색인어는 세목을 전개하기 위한 주표목의 역할을 하는 색인어들이다.

이 사이트는 앞의 성약출판사 사이트와 다르게 '가나다' 순 색인과 색인 저록 그리고 요리문답의 본문 내용 파일 등 세 개의 별도의 파일로 구성하였다. 예를 들어, <그림 7>에서 '마'를 클릭하면 그 글자의 색인 저록들의 범주로 이동하게 되고, '마'에 배열된 색인 저록 중에서 하나를 다시 클릭하면 그 색인어와 하이퍼링크된 본문으로 연결된다.



<그림 7> 하이델베르크 요리문답 웹 사이트 내 색인

<그림 7>의 '보라' 참조 기호인 '-'로 연결된 믿음의 동의어인 '참된 믿음'을 클릭하면, <그림 8>과 같이 '참된 믿음'의 세목을 포함한 완전한 형식의 색인 저록으로 이동한다. 따라서 이용자는 자신이 찾으려고 하는 세목을 선택하여 요리문답의 해당 본문으로 연결된다.

<p>참된 믿음</p> <ul style="list-style-type: none"> • <u>확실한 지식</u> • 굳은 신뢰

〈그림 8〉 페이지 내 링크로 나타난 색인 저록

〈그림 8〉 ‘확실한 지식’을 클릭하면, 아래와 같이 ‘확실한 지식’이라는 용어가 포함된 하이텔베르크 요리문답의 본문으로 연결되고, ‘참된 믿음’과 ‘확실한 지식’의 두 용어가 어떻게 연관되어 있는지를 알 수 있게 되는 것이다.

<p>21문: 참된 믿음이란 무엇입니까? 답: 참된 믿음 하나님께서 그의 말씀에서 우리에게 계시하신 모든 것이 진리라고 여기는 <u>확실한 지식</u>이며, 동시에 성신께서 복음으로써 내 마음속에 일으키신 굳은 신뢰입니다. 곧 순전히 은혜로, 오직 그리스도의 공로 때문에 하나님께서 죄 사함과 영원한 의로움과 구원을 다른 사람뿐 아니라 나에게도 주심을 믿는 것입니다.</p>

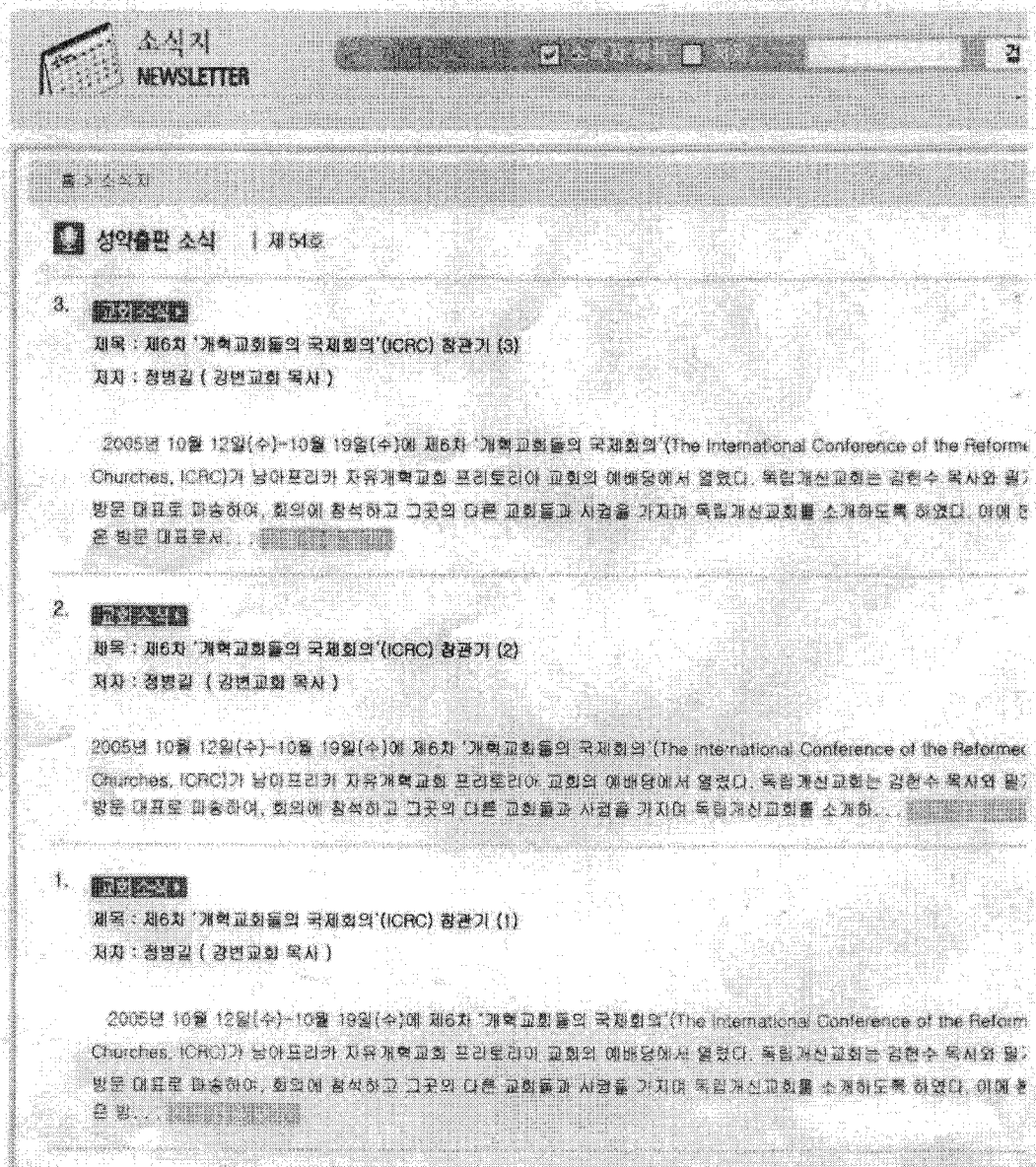
자동으로 생성한 색인어의 경우에는 색인어를 생성하는 과정보다 선정하는데 더 많은 시간이 필요로 하였다. 따라서 웹 사이트 내 색인 작성에서 표목을 자동으로 생성하는 데에는 한계가 있고, 색인자의 지적 능력이 더 중요하다는 것을 확인할 수 있었다. 그러나 표목을 자동으로 생성하는 성능이 보다 발전할 수 있다면 더 나은 웹 사이트 내 색인을 만들 수 있다는 가능성을 확인할 수 있었다.

3. 구축한 웹 사이트 내 색인에 대한 평가 및 비교

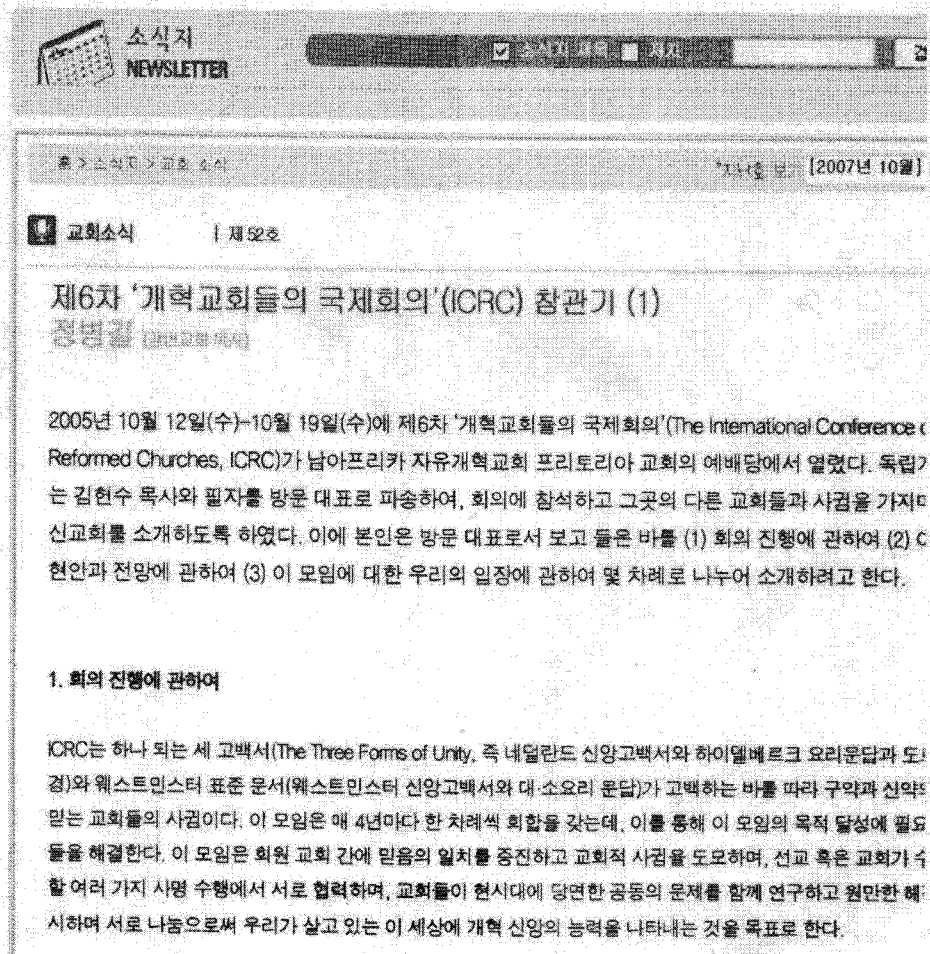
가. 성약출판사 웹 사이트 내 색인

검색 측면에서는 웹 사이트의 콘텐츠에 직접 접근할 수 있는 장점이 있고, 키워드를 모르거나 혹은 키워드를 직접 입력하지 않아도 선정된 색인어를 브라우저창으로써 이용자가 검색할 수 있다는 장점이 있다. 그러나 선정된 색인어 이외의 키워드로는 검색할 수 없으므로 사이트 내 색인이나 사이트 맵으로 보완해야 한다.

그리고 웹 사이트 내 색인에 의한 검색 결과와 사이트 내 검색에 의한 결과를 사례로 제시하면 다음과 같다. 현재 성약출판사 사이트의 메뉴 중에서 '소식지'를 대상으로 사이트 내 검색과 웹 사이트 내 색인의 결과를 '개혁교회'라는 키워드로 비교해 보았다. 사이트 내 검색 결과 3건은 <그림 9>와 같고, 웹 사이트 내 색인으로 검색한 것은 결과 중에서 하나를 제시하면 <그림 10>과 같다.



<그림 9> '개혁교회'라는 키워드로 검색한 결과 화면



<그림 10> '개혁교회' 색인어를 웹 사이트 내 색인에 의해서 검색한 화면

키워드 검색에 의한 결과는 결과 중에서 다시 선택하는 간접적인 방식이라면, 웹 사이트 내 색인은 색인어가 포함되어 있는 페이지로 바로 이동하는 직접적인 방식이라고 할 수 있다.

이와 같이 기독교 관련 웹 사이트 내 색인을 적용한 두 가지 사례에 의해서 생각할 수 있는 점은 다음과 같다. 현재 문헌정보학계나 도서관계의 현실에 비추어 볼 때 색인자가 직접 색인어를 선정해야 하는 웹 사이트 내 색인의 적용을 위해서는 색인 전문가를 양성하는 것이 필요하다. 전문 색인자가 없다는 한계로 인해 사이트 내 검색이나 사이트 맵과 비교할 때, 색인자가 직접 색인어를 선정하거나 자동 색인으로 추출된 색인어를 보정해야 하기 때문에 상대적으로 더 많은 비용이 소요될 단점을 가지고 있다.

나. 하이델베르크 요리문답 웹 사이트 내 색인

하이델베르크 요리문답은 52주 129문이 3부로 크게 나뉘고 다시 10개의 하위 주제로 구성되어 있다. 이 10개의 주제가 사이트 맵과 유사한 역할을 하고 있지만 특정 주제에서는 지나치게 문답의 내용의 길어서 내용을 검색하는 데 한계가 있다. 따라서 주단위로 접근 할 수 있는 방법과 더불어 웹 사이트 내 색인을 활용함으로써 색인어 브라우징이나 문답 내용에 대한 직접적인 접근 등에서 이용자의 검색 환경을 개선할 수 있을 것으로 보인다. 그러나 자동 색인으로 생성된 색인어들 중에서 수작업에 의해 최종적으로 선정된 색인어의 수가 예상보다 적어서, 웹 사이트 내 색인에서 자동 색인의 활용 범위를 재검토해야 할 필요성이 있음을 확인할 수 있었다.

다. 기존 웹 사이트 내 색인과 새로 구축한 웹 사이트 내 색인의 차이점

이 연구에서 구축한 웹 사이트 내 색인과 기존의 웹 사이트 내 색인과의 차이점은 아래의 <표 2>와 같다.

<표 2> 기존 웹 사이트 내 색인과 새로 구축한 웹 사이트 내 색인의 차이점

내용 \ 종류	기존 색인	구축한 색인
색인어 생성 방법	수작업	수작업과 자동 색인
색인 주제	색인 전문가	자동 색인 프로그램과 색인 전문가
구축 비용	비용 증가	비용 감소
외부 링크 수	링크 수가 다양함	링크 수가 다양하지 못함

우선 색인어 생성 방법에서는 기존 웹 사이트 내 색인은 대부분 색인 전문가에 의존하는 바가 크지만, 이 논문에서 구축한 웹 사이트 내 색인은 자동색인과 수작업을 병행하는 하이브리드 방식 이라고 할 수 있다. 따라서 당연한 결과로 색인 주제도 기존 웹 사이트 내 색인은 색인 전문가이고, 이 연구에서 구축된 웹 사이트 내 색인은 자동 색인 프로그램 즉 형태소 분석기와 색인 전문가가 함께 참여하여 한다. 그리고 구축비용은 모든 과정에서 색인 전문가가 관여하는 기존의 색인 방식 이 인건비 등으로 비용이 증가할 것이고, 상대적으로 색인 전문가가 일부분만을 관여하는 이 연구 에서 제시된 웹 사이트 내 색인은 비용이 감소할 것으로 보인다. 또한 기존 웹 사이트 내 색인은 다양한 외부 링크들을 연결하고 있었으나, 이 연구에서의 웹 사이트 내 색인은 실험의 한계로 다양 한 외부 링크를 설정하지 못했다.

V. 결 론

전통적인 인쇄 형태의 문헌이나 웹상의 문서 혹은 웹 사이트에 포함된 정보의 양이 클 때에는 이용자 자신이 찾아낸 또는 접근한 문헌이나 웹 사이트 내에서 다시 정보를 검색해야 하는 경우가 생길 수 있다. 즉 이용자가 찾고자 하는 보다 구체적인 정보를 최종적으로 획득하기 위해서는 더 직접적인 접근 수단이 필요하다는 의미이다. 이러한 상황은 전통적인 인쇄 형태의 자료뿐만 아니라 웹상에서도 발생할 수 있다.

이용자들이 직면하게 되는 어려움을 해결할 수 있는 수단으로는, 단행본과 같은 경우에는 권말 색인이고, 웹 사이트의 경우에는 사이트 색인이다. 두 가지 방법은 목록이나 분류와 같은 정보 조직의 수단과는 그 용도에서 차이가 있지만, 정보에 대한 보다 직접적인 접근책을 이용자에게 제공한다는 점에서 두 방법은 공통점이 있다. 그러나 색인의 대상이 다르고 색인 저록의 구성요소에서 차이가 나며, 색인 저록과 상세한 특정 내용을 연결하는 방식에도 차이가 있다. 가장 중요한 차이점으로는, 웹 사이트 내 색인은 소재 위치에 대한 정보를 제시할 필요가 없으며, 색인 저록과 사이트 내의 페이지나 특정한 내용을 연결하기 위해서 HTML의 <a> 태그를 활용하여 하이퍼링크를 사용하는 것이 대부분이다.

그리고 웹 사이트 내에서는 색인 이외에도 사이트 맵이나 사이트 내 검색이 유사한 역할을 수행하고는 있지만, 웹 사이트 내 색인을 작성하면 두 방법의 약점을 보완할 수 있다. 특히 웹 사이트 내 색인은 대부분 전문 색인자가 수작업으로 색인을 작성하기 때문에 자동색인으로 생성되는 색인의 단점을 보완할 수 있으며, 자동색인으로 생성된 표목만으로는 이용자의 정보 요구를 만족시킬 수 없기 때문에 2차적으로 선정 과정을 거쳐야 한다.

이 연구에서는 일반적으로 웹 사이트 내 색인을 작성할 때 사용되는 수작업 방식으로 색인을 작성하였고, 동시에 수작업 방식과 자동 색인 방식을 결합한 방법으로도 색인을 작성하였다. 그리고 기존의 웹 사이트 내 색인과 이 연구에서 구축한 웹 사이트 내 색인은 색인 주체, 색인어 생성방법, 비용, 외부 링크 설정 등에서 차이를 보였다. 이 연구에서 구축한 웹 사이트 내 색인은 색인 주체가 자동색인과 색인 전문가가 함께 참여하는 하이브리드 방식이라고 볼 수 있었고, 생성방법에서도 수작업과 자동 색인을 병행하여 활용한 면에서 기존 웹 사이트 내 색인과의 차별성을 보였다고 할 수 있다.

따라서 후속 연구로는 현재 발행되고 있는 도서에 권말 색인이 포함된 비율이 어느 정도가 되는지 혹은 그 권말 색인들이 정확한 원칙이나 지침에 의해서 제대로 작성되었는지에 대한 연구가 필요할 뿐만 아니라 기계가독목록(MARC)에서 색인어를 부여하는 수작업 색인이나 색인 작성 지침에 대한 관심이나 연구도 절실하게 필요하다. 그리고 더 본질적인 문제로 전통적으로 문헌정보학에서 사용하고 있는 정보 조직의 수단을 웹이라는 새로운 환경에서 적용하려는 연구들이 그렇게 많

지 않다는 것이다.

마지막으로 HTML의 <a> 태그를 활용하는 방법 외에도 메타데이터나 시소러스를 활용해서 웹 사이트 내의 검색 효율을 개선할 수 있는 정보 조직 수단에 대해서 지속적인 연구가 진행되어야 할 것이다. 이러한 연구에서 가장 밑바탕이 되는 것은 역시 우리가 이미 가지고 있는 방법들을 새로운 웹 환경에 적용하는 창조적인 사고일 것이다.

참 고 문 헌

- 김효열. 도서관말색인의 작성지침과 자동생성에 관한 연구. 석사학위논문, 연세대학교 대학원, 문헌정보학과. 1995.
- 독립개신교회 교육위원회 옮김. 하이텔베르크 요리문답. 서울 : 성약출판사, 2004.
- 브라운, 글렌다. 조나단 제레미. 웹사이트 색인작성. 이영자 역. 대구 : 태일출판사, 2002.
- 성약출판사. [online]. [cited 2007. 2.1] <<http://www.sybook.org>>.
- 한국도서관협회 목록위원회 편. 한국목록규칙. 제4판. 서울 : 한국도서관협회, 2003.
- 한국정보통신기술협회. 색인작성지침. 서울 : 한국정보통신기술협회, 1999.
- Browne, Glenda and Jonathan Jeremy. *Website Indexing: enhancing access to information within websites*. 2nd ed. Adelaide : Auslib Press, 2004.
- Rosenfeld, Louis and Peter Morville. *Information architecture for the World Wide Web*. Cambridge, Mass : O'Reilly, 2002.
- Leise, Fred. 2004. "Improving usability with a website index." [online]. [cited 2007.2.8]. <<http://www.web-indexing.org/article-leise.htm>>.
- American Society for Indexers. [online]. [cited 2007.2.1] <<http://www.asindexing.org/site/index.html>>.
- UNIXhelp for Users. [cited 2007.2.7] [online]. <<http://www.unixhelp.ed.ac.uk>>.
- The W3C Website Index. [online]. [cited 2007.1.20] <<http://www.w3.org/Heip/siteindex>>.
- Web Indexing SIG Brochure - United States. [online]. [cited 2007.2.8]. <<http://www.web-indexing.org/brochure.pdf>>.