

## 도서관 포털이란?



심 경

shim@irisnet.co.kr  
정보학박사  
한국도서관협회 평생회원  
(주)아이리스넷 대표

대학시절 오후 늦게 시작하는 강의를 항상 재미있게 끌어가시던 어느 교수님 생각이 난다. 하루는 “정의(定義)”란 무엇인가를 설명하는 중에 “사랑의 정의 중 가장 잘 된 것이 누구 것이냐?”는 질문을 던지셨다. 필자는 기억나는 유명한 사람 중 가장 오래 전 사람인 아리스토텔레스쯤을 머리 속으로 생각하며 “제발 저는 시키지 말아주세요”라는 마음으로 눈치를 보고 있는데, 갑자기 교수님이 “나훈아야! 사랑은 눈물의 씨앗이야!”라고 하셨던 것이다. 긴장했던 강의실은 순식간에 웃음바다가 되었고 우리는 거의 “전사(戰死)” 직전까지 가도록 웃어댔다. 사실 사랑이란 단어의 공식정의는 없다. 그렇지만 우리는 그 단어를 사용하면서 상대가 어떤 의미로 이해할까를 걱정하지 않고 말한다. 이런 개념을 미국 사람들은 “you-know concept”라고 하며, 이는 굳이 매번 의미를 설명하거나 공식정의를 외우지 않더라도 그 의미가 오해 없이 상대에게 전달되어 오해나 그 폭이 의미전달에 큰 영향을 미치지 않는 단어를 말한다.

“사랑”이란 단어만큼은 아니지만, 요즘 우리 분야의 전문용어 가운데 하나이면서 언론이나 일반인들에게까지 널리 통용되는 단어 중 하나가 “포털(portal)”이다. 그렇다면 포털이란 개념이 그렇게 모든 사람들에게 동일하거나 유사한 내용으로 이해되고 있는 개념일까? 사랑이 개인적인 관계에 한정된 개념이며<sup>1)</sup> 우리가 생활에서 듣고 배우고 느낀 개념이라면 포털은 그 반대이다. 특히 포털 제공자는 다수에

1) 물론 사랑이란 개념은 우리가 고등학교 때 배운 eros와 pathos 등 다양한 종류가 존재하지만, 여기서는 일반적으로 우리가 말하는 세속적 ‘사랑’을 의미한다.

게 서비스 제공을 위하여 포털을 구축하므로 서비스 콘텐츠, 대상 이용자 및 목적에 따라 전혀 다른 생각을 가질 수 있다. 그런데 최근 포털이란 말을 들으면 “포털 사이트의 독점이나 횡포”, “포털 상의 정보유출”이라든지, “포털의 아고라를 통한 여론형성” 등등. 어떻게 보면 긍정적이기보다 신문기사의 헤드라인으로 돌출된 부정적 이미지를 더 많이 주는 느낌이기도 하다. 그러면서 포털이라 하면 “네이버”나 “다음”과 같이 엄청나게 다양한 정보가 마치 추석날 종합선물 세트 같이 열거되어 있는 사이트를 생각하게 된다.

그 동안 도서관에서도 포털이란 말을 사용해 왔고, 그 안에는 대중이 생각하는 포털의 개념과는 분명 다른 점이 있다고 할 수 있다. 이 글에서는 포털의 출현과 초기 형태를 알아보고, 도서관 포털의 기본요소와 개발 방법에 대해서 간략히 살펴본다. 이 글을 통하여 도서관 포털 구축에 대한 해답 또는 정답을 제공할 수 있다고 생각하지 않으며, 그것이 이 글의 목적도 아니다. 다만 도서관 포털에 대하여 커다란 개념을 잡을 수 있기 바랄 뿐이다.

## 포털의 발생과 기능

원래 포털은 대기업에서 직원들이 여러 개 시스템에 분산되어 있는 회사 정보를 단일 인터페이스로 접근할 수 있도록 하기 위하여 개발되었다. 따라서 초기 포털은 대상 기업들이 자체적으로 개발하였으나 차츰 상용 포털 소프트웨어가 개발되어 제공되고 있다. 하지만 이런 상용 소프트웨어는 주로 기업을 위한 것이었으며 도서관을 위한 포털의 기능에 초점을 맞춘 것은 아니었다. 이들 포털은 일반적으로 다음 네 가지 요소를 가진다.

- 직관적이며 수정 가능한 웹 인터페이스: 한 기관의 현재 응용프로그램의 외양과 느낌에 일치되도록 디자인 될 수 있는 네비게이션이 쉬운 인터페이스
- 개인화된 콘텐츠 제공: 개인화된 콘텐츠를 제공하기 위하여 포털은 이용자 프로필 정보에 접근이 가능. 여기서 개인화는 한 사람 또는 특정 그룹을 지칭
- 보안: 이용자 프로필이 이용되며 이용자 인증은 내부 시스템에 대한 접근권한은 물론 외부 시스템에 대한 접근권한을 결정
- 커뮤니케이션과 협동: 포털은 채팅, 이메일, 공동 달력과 웹 미팅에 이용

위 특징은 일반 포털의 기능이며 이보다 다소 뒤에 출현한 도서관 포털이 위 기능 모두를 반드시 포함하는 것은 아니다. 도서관 포털은 정보접근에 역점을 두고 있으므로 채팅, 이메일 등 커뮤니케이션이나 협동 등에는 상대적으로 관심을 덜 두는 편이다.

## 도서관 포털의 발전

일반 포털과는 차별되게 도서관 포털은 정보의 통합검색에 초점을 맞추어 개발되었다. 미국의 경우 국내와는 달리 도서관 포털이 패키지 형태로 제공되고 있다. 최초의 도서관용 포털 패키지는 Lexis-Nexis에서 개발한 것으로 도서관과 법률사무소를 대상으로 개발된 것이었다. 그 뒤를 Gale, Auto-Graphics, Fretwell-Downing, MuseGlobal과 WebFeat 등이 이었다. Gale의 포털은 자신들이 제공하는 다양한 데이터베이스를 통합검색하는 형태를 취하였고, 그 이외의 회사 제품들은 요즘 우리가 생각하는 포털에 좀 더 근접한 기능을 가진 것이었다. 도서관 포털 패키지의 보다 본격적인 발전은 Endeavor, Ex Libris, Innovative Interfaces, TLC, VTLS 등 도서관자동화시스템을 공급하는 회사들이 포털 패키지를 자신들의 도서관자동화시스템의 일부로 판매하면서 시작되었다고 볼 수 있다. 이들 패키지는 도서관의 특성 상 다수 데이터베이스나 시스템의 통합검색에 치중하고 있다.

초기 도서관 포털은 다수 데이터베이스와 시스템을 통합검색하는 기능을 제공하다 보니 검색방법은 키워드 검색이 주류를 이루었고, 그 대가는 검색결과의 과잉과 부정확한 검색결과가 양산되는 것이었다. 여기서 키워드 검색이란 대상 데이터베이스나 시스템의 메타데이터 구조가 다양하므로 필드를 구분하지 않고 검색하는 것을 의미한다. 예를 들면, 인명 “심경(沈勁)”을 저자필드에서만 검색하는 것과 전체 필드에서 검색하는 것은 판이한 결과를 나타낸다. 후자의 경우 사람 이름이 아닌 “마음의 상태로서 심경(心境)”, “불교경전들인 각종 심경(心經)”, “농경법의 하나인 심경(深耕)” 등 전혀 기대하지 않았던 검색결과가 제공되는 것을 의미한다. 또한 필드를 구분하더라도 동일한 용어가 학문분야에 따라 전혀 다른 의미로 사용되는 것은 우리가 잘 알고 있는 바이다.

검색결과 과잉 문제를 줄이기 위하여 주로 사용되는 방법은 통합검색을 하되 검색대상필드를 구분하는 것이다. 그 이외에 사용되는 방법은 적합성 순위, 미리 주어진 기준에 따른 검색결과 배열 등이다. 하지만 사실 도서관 분야에서 채택하는 검색기법인 불리안 검색은 검색결과의 순위화가 이론적으로는 불가능하다<sup>2)</sup>. 그럼에도 불구하고 도서관 포털 패키지는 검색어가 검색문헌에 출현하는 빈도를 가지고 순위화하고 있다. 그러나 이 기법은 그리 효과적이지 않은 것으로 알려져 있다<sup>3)</sup>.

2) 불리안 검색은 가장 효율적인 검색모델이기는 하지만, 집합이론에 근거하여 멤버인가 아닌가를 결정하므로 질의문과 대상 문헌이 일치하는가 아닌가를 결정할 뿐, 일치하는 정도(degree)를 나타낼 수 없다. 따라서 이를 exact match 모델이라고 한다.

3) 이 기법은 Noreault, T. Koll, M. & McGill, M. J. (1977). Automatic ranked output from Boolean: searches in SIRE. JASIS, 28(1), pp. 333-339에 처음 발표되었으나 실제로는 그리 효과적이 아닌 것으로 알려진다. 일반적으로 검색어와 비교될 대상이 문헌의 제목인 경우 그 실효성이 적을 것임을 짐작하기는 어렵지 않다.

## 도서관 포털의 요소

도서관 포털은 이용자가 원하는 자료를 도서관 자체에 소장한 자료는 물론 연계를 통하여 보다 많은 대상을 검색하도록 하는 관문역할을 하므로 앞서 소개한 일반 포털의 요소와는 다소 다르다. Boss에 의하면 미국 도서관 포털 패키지가 빠르게 변화하고 있지만 단일검색 인터페이스(single-search interface), 이용자 인증(user authentication), 리소스 링킹(resource linking), 콘텐츠 확장(resource enhancement)라는 네 가지 공통 요소를 가지고 있다<sup>4)</sup>고 한다. 이들을 하나씩 살펴봄으로써 도서관 포털에 대한 방향을 알 수 있을 것이다.

### (1) 단일검색 인터페이스

단일검색 인터페이스의 필요성은 대부분의 도서관이 온라인 목록 이외에 각종 데이터베이스와 시스템, 그리고 점점 더 늘어나는 전자자원 컬렉션을 갖고 있기 때문이다. 그 동안 도서관들은 복수의 전자자원에 대해서 별도의 검색환경을 제공해 왔으나 이용자가 이들 중 어느 “제품(product)”이 자신이 찾고자 하는 것에 관련된 정보를 포함하고 있는지 알기 어려웠다. 반면 많은 대상을 하나씩 검색하는 작업은 시간이 많이 소요되고 지루한 과정이므로, 복수의 정보자원을 동시에 통합 검색할 수 있도록 하여 조사과정을 단순화 시키는 것이 필수적이었다. 이런 까닭에 Breeding은 이와 같은 “통합검색 인터페이스”가 도서관에서 요구되는 “차세대 목록(Next-Generation Catalogs)”의 중요한 요소라고도 하였다<sup>5)</sup>.

따라서 도서관 포털에서 가장 기본적인 구성요소는 복수의 전자정보원(electronic sources)에 대한 “단일 혹은 동시검색(a single or simultaneous search)”<sup>6)</sup> 인터페이스인 것이다. 이 기법은 하나의 인터페이스로 다수 대상을 검색하여 단일 인터페이스로 검색결과를 보여 주며, 그 검색결과와 출처를 밝혀준다.

이 단일검색기법은 몇 년 전 우리나라 도서관에 폭넓게 보급되었던 “메타서치”라는 시스템이 그 예이다. 이 기법은 검색대상으로 결정된 사이트의 검색방식과 검색화면을 분석하여<sup>7)</sup> “프로파일(profile)”을 구축한 후 이를 활용하여 통합검색을 한다. 그러므로 도서관 포털 패키지는 Z39.50, HTTP, SQL, XML Web Services 등 프로토콜(protocol)이라 불리는 가능한 모든 표준

4) Boss, R. W. (2002). How to plan and implement a library portal. *Library Technology Reports*, 33(6).

5) Breeding, M. (2007). Next generation library catalogs. *Library Technology Reports*, 43(3).

6) 이는 여러 가지 이름으로 불리며, 통합검색(federated search), 브로드 캐스트 검색(broadcast search), 메타검색(meta search), 교차검색(cross-searching) 등이 있다.

7) 통합검색의 기본은 우선 어플리케이션과 검색되는 각각의 타깃 사이에서 일어나는 “배후(behind-the scene)”의 커뮤니케이션에 있는데 이를 분석하는 것이다.

검색방식은 물론 사이트 고유의 방식도 지원해야 하며, MARC, 더블린 코어, EAD, TEI, XML 등 각종 메타데이터 포맷도 지원이 가능해야 한다.

다시 말하여 우리는 Z39.50가 복수 도서관시스템을 하나의 인터페이스로 검색 가능하도록 하는 것을 알고 있다. 그런데 이를 지원하지 않는 곳은 대상 사이트가 사용하는 검색방식을 분석하여 응용프로그램에 통합해 놓음으로써 이용자 검색은 Z-프로토콜 검색처럼 단일 인터페이스로 제공되는 것이다. 따라서 통합검색 패키지는 광범한 타깃 자원을 포함하기 위하여 앞서 언급한 많은 “탐색 및 검색(search and retrieval)” 방법을 지원할 수 있어야 한다

이 통합검색 시스템은 일반적으로 타깃 자원에서 보내온 결과들의 디스플레이를 조정하는 기능을 가진다. 그 중 하나로 다양한 자원에서 가져온 레코드들은 보통 XML과 같은 표준구조로 저장되고, 조직되고, 정렬되어 표준포맷으로 이용자에게 제시된다. 이때 중복레코드를 대상으로 중복제거(deduplication)가 이루어지기도 한다. 또한 이 시스템은 모든 검색 타깃으로부터 결과를 통합 정렬하거나 출처에 따라 분류된 상태를 그대로 보존할 수도 있다.

일반적으로 외국은 통합검색 구현업체가 타깃 사이트와 협약을 맺는 경우가 많으나 국내의 경우는 일방적으로 프로파일을 구축하여 서비스함으로써 상대 사이트에 변화가 발생하면 검색 결과 화면에 오류가 발생할 확률이 높다. 그러므로 국내에서는 검색결과를 통합 정렬된 형태로 제공하는 것보다 사이트 별로 보존하여 제공하는 것이 더 안정적일 것이다.

## (2) 이용자 인증

이용자 인증은 이용자가 도서관의 데이터베이스나 자원을 이용할 권한이 있는지를 결정하는 절차로 우리에게 전혀 낯설지 않다. 이 인증은 보통 프록시 서버에서 이루어지며, 이용자가 이용할 권한을 가진 자원에만 접근할 수 있게 제한하는 목적을 가진다. 보통 모든 도서관은 이용자로서 하여금 웹 상에서 도서관의 OPAC이나 도서관이 제공하는 일반적인 정보는 이용하도록 하는 것이 원칙이므로 이용자 인증은 도서관 자료에 대한 접근보다는 대부분 도서관이 구독하는 데이터베이스(subscription databases) 혹은 라이선스 제한을 갖는 그 밖의 자원들 때문에 인증 절차를 적용하는 경우가 대부분이다.

## (3) 리소스 링킹

리소스 링킹은 도서관으로 하여금 전자자원들이 매끄럽게 서로 연결되도록 허용하는 것이다. 예를 들어, 색인이나 초록이 원문 데이터베이스에 링크될 수도 있고, 자관의 서지 레코드들이 리뷰나 전자책에 링크될 수도 있다. 초기 도서관 포털 패키지는 이 시스템 구현을 위하여 패키지를 수용한 도서관이 전자자원으로의 연계를 자체적으로 생성하도록 요구하였으나,

Endeavor Information Systems를 시작으로 벤더와의 라이선싱 소프트웨어와 연계정보를 미리 구축한 데이터베이스가 함께 제공되고 있다. 하지만 도서관이 원하는 모든 리소스로서의 링킹이 벤더들에 의하여 제공될 수는 없으므로 도서관 측이 다소의 수작업을 감수해야 한다. 리소스 링킹의 표준적이며 가장 발전된 형태는 이미 이전 글에서 살펴 본 링킹시스템으로 현재는 학술분야에 주로 채택되고 있으나 이 방법은 조만간 그 적용범위가 확장될 것으로 기대된다.

#### (4) 콘텐츠 확장

콘텐츠 확장(content enhancement)은 도서관에서 단순히 MARC 기반의 서지레코드를 제공하는 것을 넘어 자료의 목차, 표지 이미지, 저자의 전기정보, 리뷰 등과 같은 정보에 대한 링크를 제공함으로써 서지정보를 확장하는 것을 의미한다(그림 1과 2 참조). Breeding(2007)도 통합도서관자동화시스템 OPAC 모듈이 제공하는 대표적인 확장기능으로서 이 콘텐츠 확장을 든 바 있다. 이는 어찌 보면 웹 2.0의 기술 중 하나인 매쉬 업(mash-up)에 가깝다고도 할 수 있다. 특히 저자의 전기정보는 인명사전과의 연계로 실현할 수 있기 때문이다.

콘텐츠 확장을 통해 제공되는 정보는 도서관에서 자체적으로 구축할 수도 있겠으나 이를 전문으로 하는 회사가 구독 기반(subscription basis)으로 제공하는 것이 일반적이다. 미국을 비롯한 유럽지역의 주요한 도서관들이 사용하는 상용 통합 도서관자동화 시스템들을 비교 평가한 보고서에 따르면, 대부분의 공공, 대학 및 학교도서관들이 사용하고 있는 Aleph, Horizon, Library Solution, Millennium, Polaris, Unicorn, Voyager 등의 시스템이 모두 Syndetic Solutions의

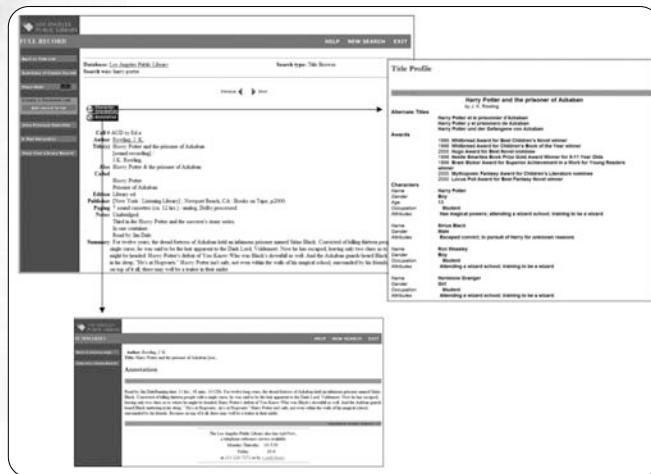


그림 1. LA 공공도서관 레코드의 '주인공 정보'와 '요약'의 콘텐츠 확장



그림 2. LA 공공도서관 레코드의 '표지 이미지'의 콘텐츠 확장

콘텐츠를 이용하고 있다<sup>8)</sup>. 국내에도 이와 유사한 서비스를 제공하는 업체가 존재한다. 이러한 업체를 선정 시 고려할 점은 제공 데이터의 저작권이 해결되었는가를 확인하는 것이다.

이 밖에도 도서관 포털 소프트웨어가 제공하는 인터랙티브 기능은 이메일, 채팅, 포럼과 여론조사(poll) 등이 있으며 일부 도서관들은 인터랙티브 서비스를 통해 이용자 개인의 대출상태, 벌금, 대기(holds) 및 선호하는 특정한 사이트, 최근 검색전략의 이력 등 저장된 정보를 제공하기도 하고, 저장된 프로파일에 따라 이용자에게 정보를 푸시(push)하기도 한다. 이러한 기능들의 대부분은 국내 도서관자동화시스템이 오래 전부터 제공하던 것으로 그리 새로운 것은 없으니, 우리는 이미 어느 정도 도서관 포털의 기능을 가지고 있는 셈이다.

## 결 언

이 글에서는 도서관 포털의 요소와 경향을 외국문헌을 중심으로 살펴보았다. 외국의 사례를 국내 도서관 포털 또는 자신 도서관 포털과 비교를 원하는 독자들은 상용 포털 제품인 Auto-Graphics Inc.의 Agent로 구축한 캐나다의 포트무디 공공도서관<sup>9)</sup>, Fretwell-Downing Informatics의 ZPortal로 구축한 영국의 국립보건도서관<sup>10)</sup>, 또는 자체적으로 포털을 구축한

8) Breeding, M. (2004). Integrated library software: a guide to multiuser, multifunction systems. Library Technology Reports, 40(1)

9) <http://library.portmoody.ca/default.htm>

10) <http://www.library.nhs.uk/Default.aspx>

호주의 Libraries Australia, 미국 캘리포니아 주의 California Digital Library 등의 사이트를 방문해 볼 것을 권한다<sup>11)</sup>.

지금까지 국내 도서관계에서도 포털이란 말이 종종 사용되어 왔지만, 특히 최근 국립중앙도서관이 “디지털 지식정보 포털 협력망”, “다문화 지식정보 포털 협력망”, “지역 지식정보 포털 협력망” 및 “정책정보 포털 협력망” 등 다양한 포털 협력망을 구축하겠다고 하면서<sup>12)</sup> 더 그 용어가 부각되는 인상을 준다.

도서관 포털의 주요 기능은 이용자에게 원하는 정보를 편리한 방법으로 제공하는 것이라고 생각한다. 필자는 도서관이 아닌 국내 다른 기관의 포털을 방문했다가 너무 많은 것들이 한꺼번에 “우겨 넣어진” 복잡함에 당황하고 찾을 것을 못 찾은 채 떠난 적이 많이 있다. 도서관 포털은 이용자에게 보다 많은 접근기회를 제공하기 위하여 많은 링크와 데이터베이스를 제공하는 것도 중요하겠지만 그들을 이용자가 얼마나 쉽게 발견하고 이용할 수 있는가가 더 중요한 사안이라고 생각한다. 과거 국내 도서관 포털이 전자를 중요시 하였다면 후자에 노력을 기울일 때다. 문득 수년 전 읽은 책의 제목이 생각난다. “Don’t Make Me Think! A Common sense Approach to Web Usability”라는 책이다. 도서관 포털은 연구대상이 아니다. 이용자가 들어와서 직관적으로 원하는 것을 얻도록 이해하기 쉽고 이용하기 편해야 하는 것이라는 점에서 우리는 심심할 겨를이 없다. 이용자가 생각하지 않게 하기 위해서 우리는 엄청나게 많은 생각을 해야 하기 때문이다.

마지막으로 콘텐츠 확장과 관련하여 이는 단일 도서관이 수행하기에는 감당하기 어려울 만큼 많은 노력과 경비가 소요되어 현실성이 없으며, 그렇다고 국내 도서관자동화 업체가 미국 업체처럼 콘텐츠까지 판매하는 추세를 따라가기에는 영세하여 근시일 내에 실현되기 어렵다는 판단이다. 최근 사서들이 구성한 협의체 모임에 우연히 참석할 기회가 있었는데 거기서 “열정”이란 단어의 의미를 처음 느껴 보았다. 이와 같이 열정을 가진 사서들의 협의체에서 이러한 사업을 한 부분으로 수행하면 좋지 않을까 하는 생각을 해 본다.

11) 이러한 참고 사이트를 일일이 방문하여 분석할 시간이 없는 독자는 필자에게 요청하면 정리된 파일을 제공할 수 있다.

12) 국립중앙도서관 (2008). “2008년 선진일류 도서관 발전 세미나” 자료집. (2008년 6월 26일, 국립중앙도서관 국제회의장)