

디지털도서관의 통합검색 방식에 관한 연구*

A Study on the Methods of Integrated Search in Digital Libraries Environment

이 수 상(Soo-Sang Lee)**

< 목 차 >

- | | |
|---------------------|--------------------|
| I. 서론 | 1. 도서관 포털과 통합검색 |
| II. 디지털도서관과 통합 | 2. 도서관 포털의 참조사례 분석 |
| 1. 디지털도서관의 특성 | 3. 통합검색 방식의 유형 분석 |
| 2. 디지털도서관의 통합 영역 | IV. 결 론 |
| III. 디지털도서관 통합검색 방식 | |

초 록

이 연구는 디지털도서관에서 제공 가능한 통합검색 방식의 유형과 특성을 분석할 목적으로 다음의 세 가지 작업을 수행하였다. 첫째, 이용자 관점에서 디지털도서관의 발전단계와 일반적인 통합 방식의 특성을 정리하였다. 둘째, 현재 국내외 대표적인 도서관 포털의 모범사례인 영국의 JISC IE, 미국의 NSDL의 OCKHAM, 그리고 한국의 국가지식포털을 대상으로 통합검색의 현황과 특성에 대하여 검토하였다. 셋째, 디지털도서관에 영역에서 많이 적용되고 있는 통합검색 방식의 유형을 메타통합검색과 분산검색으로 구분하여 각각의 기법적 특성을 도출하였다.

주제어: 디지털도서관, 통합검색, JISC IE, OCKHAM, 국가지식포털, 메타통합검색, 분산검색

ABSTRACT

This study intends to analyze the integrated search methods in digital libraries environment. To categorize various integration methods, I investigated the developmental properties and integration types of digital libraries from a users point of view. And then, I reviewed the current best practices of library portal which are JISC Information Environment, NSDL OCKHAM and Korea Knowledge Portal initiatives. Nextly, I derived the two methods of integrated search mechanisms such as aggregated search and distributed search that is useful in the digital libraries environment.

Key Words: Digital Library, OCKHAM, JISC IE, Korea Knowledge Portal, Integrated Search, Aggregated Search, Federated Search

* 이 논문은 부산대학교 자유과제 학술연구비(2년)에 의하여 연구되었음.

** 부산대학교 문헌정보학과 조교수(sslee@pusan.ac.kr)

• 접수일: 2006년 5월 29일 • 최초심사일: 2006년 6월 1일 • 최종심사일: 2006년 6월 22일

I. 서론

디지털도서관은 디지털장서를 대상으로 각종 정보서비스를 제공하는 도서관이다. 전통적 도서관이 주로 아날로그 자원들을 대상으로 장서를 구성하고 이용자에게 각종 정보서비스를 제공하는 것이었다면, 디지털도서관은 대상 장서가 디지털자원이라는 것과 제공되는 정보서비스의 유형이나 방식에 있어 디지털적인 특성을 나타낸다. 그러나 도서관 서비스로서 적용되는 기본원칙에는 별다를 차이가 없다고 할 수 있다. 이러한 관점은 온라인 무료 백과사전인 '위키피디아(Wikipedia)'의 정의¹⁾를 통해 확인이 가능하다. 여기서, 디지털도서관은 컴퓨터라는 수단을 통해 접근이 가능한 디지털자원을 주된 대상으로 하는 도서관이라고 명확히 정의하고 있기 때문이다. 요약하면, 디지털도서관은 소장 및 접근 형태의 디지털자원들을 대상으로 도서관의 정보서비스 기능을 수행하는 시스템인 것이다.

그동안 논의된 디지털도서관의 정의들을 일일이 설명하는 것은 어려운 일이다. 1990년대부터 본격적으로 등장하여 현재 겨우 15년 정도의 역사를 가진 실체라고 하기에는 믿기지 않을 정도로 다양한 유형의 디지털도서관이 생겨났고, 각각은 자신의 실체와 역할에 걸맞은 서비스를 하고 있기 때문이다. 또한 시기별로 다른 키워드를 사용하면서 역할과 기능이 강조되어 왔다. 물론 이들에게서 공유되는 공통점이 없는 것은 아니기에, 그것을 중심으로 일리가 있고 사람들로부터 공감을 살 만한 정의를 구성할 수 있다. 중요한 사실은 디지털도서관도 일반적인 도서관과 마찬가지로 유기체적인 특성을 가지고 있다는 점이다. 디지털도서관도 주변의 환경 특히 정보와 기술적 환경의 변화, 이용자 정보요구 등에 따라 진화하고 발전하기 때문이다.

아무튼 최근의 여러 연구문헌들을 살펴보면, 디지털도서관의 특성 중 '통합성', '개방성', 그리고 '보존성'의 3가지에 대한 논의가 많은 것 같다.²⁾ 미국의 DLF, 영국의 JISC 등에서 디지털도서관의 연구개발에 집중적으로 투자하고 있는 핵심 키워드가 바로 이러한 특성들과 밀접한 관련이 있기 때문이다. 보존성은 디지털도서관의 핵심자원인 디지털장서의 영구적 보존과 접근을 제공하는 즉, 디지털 아카이브³⁾적인 역할을 강조하는 것으로 최근 들어 중요성이 많이 부각되고 있다. 보존

1) http://en.wikipedia.org/wiki/Digital_library. (인용 2006. 04. 10).

2) 알리 쉬리(Ali Shiri)는 2003년도 논문을 통해 디지털도서관 연구의 최근 동향에 대하여 정리한 바 있는데, 주요 내용을 살펴보면 다음과 같다. 첫째, 디지털도서관의 이론적 틀을 구성하는 3가지 요소는 사람, 정보자원 그리고 기술이며, 최근의 연구는 주로 이 3가지 영역에 관한 과제들이다. 둘째, 최근 디지털도서관에 관한 주요 현안범주를 8가지로 정리하고 있다: 1) 아키텍처, 시스템, 도구, 기술, 2) 디지털 콘텐츠와 장서, 3) 메타데이터, 4) 상호운용성, 5) 표준화, 6) 지식조직화 시스템, 7) 이용자와 사용성, 8) 법률, 조직, 경제, 사회적 이슈와 같은 거버넌스 문제. Ali Shiri, "Digital library research: current developments and trends," *Library Review*, Vol.52, No.5(2003), pp.198-202.

3) 디지털도서관과 디지털 아카이브는 디지털 자원을 관리하고 유용한 서비스를 제공하는 시스템이라는 관점에서 보면, 유사하고 서로 공유되는 역할과 기능을 가지고 있다. 전통적인 도서관과 기록관(아카이브)의 차이처럼, 이들을 구분한다면 두 가지 특성에 대해 논의가 가능하다. 첫째, 가장 중심이 되는 기능에 관한 것으로 전자는 '검색', 후자는 '보존'이라 할 수 있다. 그러나 디지털도서관에서 '보존' 기능이, 디지털 아카이브에서 '검색' 기능이 중요하게 되어 이 관점에서의 구분은 애매해 진다. 둘째, 관리하고 서비스하는 대상 자원의 관리 단위에 관한

성과는 달리, 통합성과 개방성은 다양한 유형의 디지털도서관에 공통적으로 적용되며, 또한 비슷한 역할을 하는 다른 정보시스템과는 뚜렷하게 구분하는 특성이라 할 수 있다

본 연구에서는 이 특성들 중에서 디지털도서관의 통합성에 초점을 맞추고 있다. 디지털도서관의 통합성 문제는 다양한 관점과 수준에서 분석하고 설명해야 한다. 각종 시스템의 통합, 정보자원의 통합, 서비스의 통합, 이용자 통합, 보안 통합 등 대상 영역을 고려해야 하기 때문이다. 따라서 여기서는 정보접근 환경의 통합 즉, 통합검색 문제를 중심으로 관련된 문헌들과 국내외의 주요 모범적인 참조사례들을 통해 디지털도서관의 통합성을 구성하는 수준이나 방식들을 논의하게 된다. 통합검색은 디지털도서관 통합성 문제의 기본이자 핵심이다. 그러므로 디지털도서관의 통합검색의 기본적인 원칙과 관련된 방식을 유형화 하고, 이를 지원하는 기법들과 특징들을 분석할 것이다.

II. 디지털도서관과 통합

1. 디지털도서관의 특성

디지털도서관의 통합 문제를 논하기에 앞서, 디지털도서관의 성격을 우선 정리할 필요가 있다. 왜냐하면 성격에 따라 통합의 대상과 방법을 구분할 수 있기 때문이다. 그러나 서론에서 언급하였듯이 디지털도서관에 대한 성격을 규정하는 것이 쉬운 일은 아니다. 그래서 통합문제에 도움이 될 만한 몇 가지의 성격정의를 살펴볼 것이다. 먼저, '위키피디아' 사전에서는 디지털도서관을 아주 의미 있게 정의하고 있다. 디지털도서관은 디지털장서, 서비스, 그리고 인프라로 구성되며, 이것은 평생교육, 연구, 학술유통과 보존 활동을 지원하며, 정보의 민주화와 관련된 과정이라고 하기 때문이다.⁴⁾ 이것은 디지털도서관의 구성요소와 용도를 중심으로 설명한 것이며, 특히 제시된 3가지 구성요소인 디지털장서, 정보서비스, 그리고 플랫폼으로서의 인프라는 통합을 논의할 때 중요한 키워드이다.

선마이크로시스템은 2002년에 "디지털도서관의 기술동향"이라는 간단한 백서를 발간한 바 있는데, 여기서는 디지털도서관을 전통적인 도서관에서 이용자들이 수행하던 기능과 자원이 전자적으로 확장한 것이라 정의하고 있다. 즉, 전자적 형태의 정보자원을 대상으로 하며, 그것은 멀티미디어 저장소에 저장되며, 웹기반의 서비스를 통해 이용가능하다고 한다.⁵⁾ 디지털도서관은 저장된 정보

것이다. 디지털 아카이브는 주로 유일본의 원자료(primary sources)를 대상으로 하는 시스템이며, 디지털도서관은 연구자, 교수 및 교사, 그리고 학생의 연구/강의/학습 등과 관련된 양질의 정보자원을 대상으로 하는 시스템이다. 그러나 전자는 개별자원보다 장서 단위에서 관리활동을 수행하며, 후자는 장서보다는 개별자원(아이템) 단위에서 관리활동을 한다는 점에서 기본적인 차이가 있다. 그렇지만 최근 디지털도서관은 장서 기반의 관리를 모색하고 있으며, 디지털 아카이브에서도 개별자원의 관리 문제에 관심을 나타내고 있다. 그러기에 대상 자원의 관리 단위에 대한 구분도 명확하지 않다고 할 수 있다.

4) http://en.wikipedia.org/wiki/Digital_library. (인용 2006. 03. 15).

자원을 대상으로 웹기반의 통합정보 서비스를 제공하는 전자적 환경이라는 것이다.

한편, 미스초(William H. Mischo)는 디지털도서관의 핵심은 게이트웨이나 네비게이션 서비스가 가능한 인터페이스를 제공하고, 분산통합 방법을 사용하여 이용자들에게 다양하고 분산된 정보 환경을 제공하는 것이어야 한다고 주장하고 있다.⁶⁾ 즉, 디지털도서관은 다양하고 분산된 디지털장서들을 대상으로 통합검색을 제공하는 것을 목적으로 규정하고 있다.

이와 같은 성격 정의에서 나타나는 공통적인 특성과 키워드는 다양하고 분산된 디지털장서를 통합하고, 통합검색과 같은 통합서비스를 제공하는 통합플랫폼이라는 점이다. 디지털장서의 통합, 정보서비스의 통합, 그리고 통합 플랫폼의 특성으로 디지털도서관을 정의하고 있다. 즉, 디지털도서관은 다양한 유형의 대규모 정보자원을 구축하여 이용자에게 통합정보 서비스를 제공하는 정보 플랫폼인 셈이다. 그러므로 디지털도서관은 내외부에 존재하는 다양한 정보자원의 디지털장서들을 통합하며, 이를 대상으로 통합검색과 같은 부가가치 서비스를 제공하며, 이와 같은 디지털장서의 관리와 통합검색 서비스를 위한 각종 기능들을 수행하는 컴포넌트들이나 시스템들이 통합된 디지털 플랫폼이다.

사실, 디지털도서관은 전자 또는 디지털 환경에 따라 도서관의 각종 서비스 체계를 변화시킨 것으로 정보환경의 변화에 따른 발전 단계는 <표 1>과 같이 구분할 수 있다. 각각의 단계가 확연하게 구분되는 것은 아니지만, 몇 가지 주요한 특성들을 고려하여 개념적이고 논리적으로 정리한 것이다. 3가지 유형의 시스템들은 개별적인 시스템으로 운영되기도 하며, 하나의 시스템 패키지로 운영되기도 한다. 그래서 디지털도서관은 모든 단계의 시스템 유형을 포괄하기도 하고, 개별적인 역할을 하는 독립적인 시스템으로 존재하면서, 필요한 수준에서 디지털도서관과 통합하고 연계할 수 있다.

<표 1> 디지털도서관의 발전 단계

	단계-1	단계-2	단계-3
시스템 구분	도서관자동화시스템 (Library Automation System)	전자도서관(Electronic Library)	디지털도서관(Digital Library)
정보자원	아날로그 자원(자원 중심)	전자 자원(자원 중심)	하이브리드 자원(장서 + 자원 중심)
주요 역할	소유	소유 + 접근	소유 + 접근 + 보존
핵심시스템	목록시스템	콘텐츠관리시스템(eCMS)	장서관리시스템(dCMS)
주요 메타데이터	서지 메타데이터(MARC 21, KORMARC 등)	자원 메타데이터(DC, EAD, MODS, VRA, METS 등)	장서 메타데이터(RSLP CD, DC CD AP 등)
검색 인터페이스	OPAC	Web OPAC 도서관 웹사이트	도서관 포털
검색방식	개별 검색	개별 검색	통합 검색
기타			Global DL Open DL

5) Sun Microsystem Inc., "Digital Library Technology Trends," 2002.

www.sun.com/products-n-solutions/edu/whitepapers/pdf/digital_library_trends.pdf(인용: 2006. 3. 15).

6) William H. Mischo, "Digital Libraries: Challenges and Influential Work," *D-Lib Magazine* 11(7-8), 2005. <http://www.dlib.org/dlib/july05/mischo/07mischo.html>(인용 2006. 3. 15).

컴퓨터 시스템의 발전은 도서관 운영 전반에 대한 자동화의 요구를 발생시켰고, 오프라인 목록 카드를 MARC이라는 기계가독형의 온라인 목록으로 전환되는 등의 변화를 촉진시켰다. 이 때 등장한 것이 바로 도서관자동화시스템이다. 목록시스템을 핵심으로 하여 수서, 열람 등 업무 영역의 운영에 있어 자동화가 가능해 졌다. 또한 이용자는 OPAC을 제공하는 검색시스템을 통해 각종 정보검색을 수행하게 되었다.

한편, 전자도서관은 각종 전자자원 콘텐츠의 효율적인 관리의 필요성에 따라 등장한다. 각종 유형의 전자자원 콘텐츠의 관리를 위하여 다양한 메타데이터를 사용하는 콘텐츠관리시스템을 근간으로 한다. 또한 Web OPAC이나 도서관 웹사이트와 같이 웹기반의 접근방식을 이용자에게 제공한다. 전자자원은 도서관이 직접 소장하는 것을 중심으로 하며, 외부소장의 디지털장서도 제공한다.

디지털도서관은 앞의 두 시스템과 달리 장서와 자원 모두를 중심으로 하며, 정보자원의 소유와 접근 환경뿐만 아니라 보존의 역할도 수행하는 등 그 역할이 확대된 시스템이다. 전자도서관과 디지털도서관은 전자자원을 관리하는 시스템의 유형이지만, 후자는 전자보다 규모나 역할, 관련된 표준화와 요소기술에서 차이가 난다. 따라서 디지털도서관은 관리하는 장서의 규모가 크고 관련된 각종 서비스의 속성이 다양하므로 이를 효율적으로 관리하는 장서관리시스템(Digital Collection Management System)을 근간으로 하여 각종 유관한 시스템이나 요소기술들을 포함한다. 이용자는 도서관 포털이라는 인터페이스로 접근하며, 내외부의 모든 장서들을 대상으로 하는 통합검색을 원칙으로 한다.

2. 디지털도서관의 통합 영역

디지털도서관은 기본적으로 '통합'이라는 개념과 방법론을 전제로 정보서비스를 구성한다고 할 수 있다. 디지털도서관은 내외부에 존재하는 각종 디지털 정보자원들을 효율적으로 관리하고, 또한 이들로부터 효율적인 정보자원의 검색(또는 탐색)⁷⁾ 서비스를 제공하는 역할을 수행하게 된다. 디지털 정보자원의 관리와 검색이라는 역할에서 요구되는 효율성은 통합을 통해 극대화될 수 있다. 즉, 디지털도서관의 효율성은 다양하고 방대한 규모의 정보자원들을 통합환경에서 관리⁸⁾하고, 이

7) 검색(search)과 탐색(discovery)은 둘 다 디지털도서관이나 정보검색 시스템에서 이용자가 정보자원을 찾아내는 행위를 말한다. 굳이 구분한다면, 전자는 이용자가 알고 있는 정보원(데이터베이스)을 대상으로 검색하고자 하는 경우에 사용하는 개념이며, 후자는 이용자가 규모와 내용을 잘 모르는 정보원으로부터 정보자원을 발견하는 경우(resource discovery)에 사용하는 개념이다. 디지털도서관이나 인터넷 검색포털과 같이 방대한 규모의 분산자원을 대상으로 하는 경우 분산검색의 의미가 있는 탐색이 더 어울린다고 할 수 있다. 검색/탐색 행위에 의해 식별된 정보자원은 접근(access) 행위에 의해 내용의 확인이 가능하고, 이것은 제공(delivery) 행위에 의해 이용자에게 전달된다.

8) 정보자원의 통합관리는 콘텐츠 관리시스템이나 장서 관리시스템 차원에서 구현된다. 전자는 디지털콘텐츠라고

를 기반으로 검색 행위를 제공할 경우 높아진다는 것이다.

디지털도서관에서 통합은 대체로 세 가지 영역으로 구분할 수 있다. 우선 정보자원 수준에서의 통합, 검색과 같은 정보서비스 수준에서의 통합이 가능하다. 전자는 정보자원(또는 장서) 관리자의 입장에서 본 통합이며, 다양한 유형의 정보자원이나 장서의 통합을 의미한다. 후자는 정보서비스 이용자의 입장에서 본 통합으로 정보자원으로부터 제공되거나 연관되는 각종 정보서비스의 통합이다. 한편, 디지털도서관을 구성하는 각종 응용시스템(이용자인증, 보안, 도서관자동화, 전거/시소러스, 식별체계, 링킹 등과 같은 응용기능을 수행하는 시스템)을 통합하여 하나의 거대한 통합플랫폼을 제공하는 것도 통합의 유형이 된다. 이와 같이 각종 정보인프라의 통합은 시스템 관리자의 입장에서 본 통합이다.

궁극적으로 디지털도서관은 통합정보 환경(Integrated Information Environment)을 구축하여 이용자에게 통합 서비스를 제공하는데 있다. 디지털도서관의 통합정보 환경이 어떠한 통합방식을 통해 구현되든지 간에, 이용자에서 보여 지는 것은 통합정보 서비스일 것이다. 한편, 도서관이 제공하고 있는 각종 서비스들 가운데 가장 중요한 것은 단연 '검색' 서비스이다.⁹⁾ 이용자들이 원하는 자료의 소장여부 확인에서부터 서지정보의 확인, 실질적인 자료의 입수에 이르는 과정으로 구성되는 정보검색은 도서관 서비스의 핵심이라는 말이다. 만일, 도서관 내에 존재하는 시스템들이 독자적인 검색환경을 제공하거나 각 시스템들간 상호연용성이 부족한 상황에서 검색 서비스를 제공하는 경우, 이용자는 불편하고 비효율적이라고 느끼게 된다. 디지털도서관의 '통합검색'은 이러한 문제점을 해결하는 방안으로 요구된다.

따라서 여기서는 디지털도서관 정보서비스의 통합 영역 중에서 특히 이용자 중심의 통합검색 서비스에 집중하여, 디지털도서관에서 통합검색은 어떠한 방식과 세부적인 기법으로 실현되는지를 규명해보기로 한다.

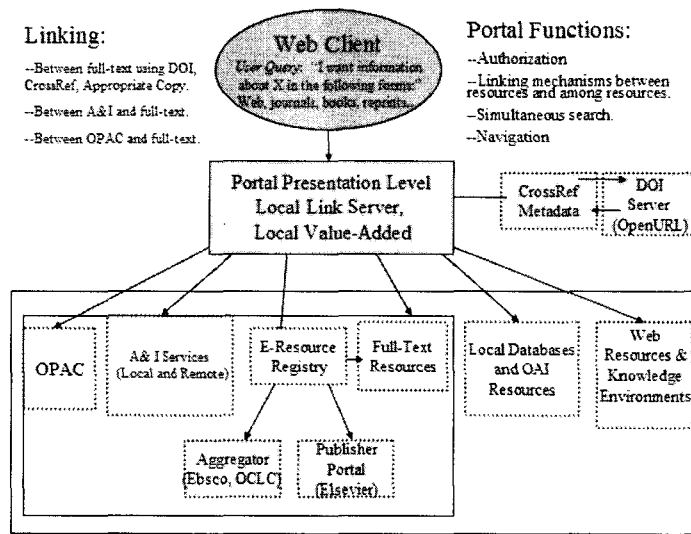
하는 개별정보자원 수준에서 그것의 수집, 정리, 유지 등에 해당되는 기능을 제공하는 것이라면, 후자는 개별정보 자원의 집합체인 디지털장서 수준에서 통합관리에 필요한 기능들을 제공하게 된다. 콘텐츠 관리시스템은 기관 내부에 소장하고 있는 디지털자원에 대한 관리기능에 초점을 맞춘 것이라면, 장서 관리시스템은 내외부에 존재하는 다양한 디지털장서의 관리기능에 초점을 맞춘 것이다. 각각은 전자도서관과 디지털도서관의 핵심 시스템이라고 하지만, 디지털도서관의 입장에서 보면 이 두 가지 시스템 모두가 필요하다. 한편, DLF에서는 전자자원의 라이프사이클에 기반하는 전자자원 관리시스템(electronic resources management system)의 기능적 요건을 제시하기도 하였다.

9) 검색 서비스 이외에도 도서관에서는 다양한 정보서비스들을 제공하고 있다. 도서관에서 이용자들에게 일방적으로 제공하는 서비스와 이용자들이 도서관에 요청하는 서비스로 크게 구분이 가능하다. 전자는 이메일이나 전화를 통한 도서반납요청, 대출내역조회, SDI 서비스, 지정도서 안내 등이 있고, 후자는 희망도서신청, 예약신청, 상호대차(원문복사), 참고봉사(질의) 등이 있다. 이와 같은 정보서비스에 의해 제공되는 정보는 주로 연성정보이며, 검색서비스에 의해 접근하고 제공되는 정보는 주로 경성정보의 형태이다.

Ⅲ. 디지털도서관의 통합검색 방식

1. 도서관 포털과 통합검색

오늘날 연구/강의/학습 영역에 있어 가치 있는 정보는 상업적인 정보, 전문학회나 기관의 정보, 개방접근 사이트 등 분산환경에서 제공되고 있기 때문에 효율적인 검색과 원문제공 등과 같은 각종 정보서비스를 수행하는 일은 쉽지 않다. 그러므로 통합환경의 기술적 아키텍처를 기반으로 각종 표준화와 요소기술들을 적용한 포털환경에서 통합검색 서비스를 제공하게 되는데, 대개는 <그림 1>처럼 구성될 수 있다.¹⁰⁾



<그림 1> 통합검색 다이어그램

이 그림에서 알 수 있듯이 도서관 포털의 인터페이스에서는 OPAC, 색인초록 DB, 전자자원 등 록기(Ebsco, OCLC, Elsevier의 포털), 원문자원, 자관의 DB나 OAI 자원, 웹자원 및 지식환경과 같은 정보원들을 대상으로 이용자에게 통합검색 서비스를 제공하고 있다. 제공되는 도서관 포털의 기능으로 이용자인증, 링킹기술, 동시검색, 네비게이션 등을 제시하고 있지만, 이 모든 것은 통합검색과 관련한 기능들이다.

궁극적으로 도서관에서의 통합검색 아키텍처는 이 그림에서 제시하는 표준화와 요소기술을 준 용하는 방식으로 구성되어야 한다. 그러므로 디지털도서관의 통합검색은 도서관 포털에서 다양한

10) William H. Mischo, 앞의 논문.

표준화와 요소기술을 기반으로 제공된다. 통합검색환경으로서 도서관 포털은 이용자에게 디지털도서관에서 소장하고 있는 정보자원과 디지털도서관에서 접근 가능한 외부 정보자원을 아주 편리한 방식으로 접근을 할 수 있는 이용자 인터페이스이다.

콕스(Andrew Cox)는 이러한 통합 도서관 포털에서 제공 가능한 서비스로 검색과 브라우징이 가능한 전자자원의 데이터베이스, 서지 데이터베이스나 다른 메타데이터의 데이터베이스의 교차검색, 단일의 공통 인터페이스, 이용자 인증, 이용자 집단의 프로파일 제공, 이용자의 선택에 따른 커스텀마이징(개인화), 각종 보고서 작성 등을 예시하고 있다.¹¹⁾ 다비스(Ron Davies)는 오늘날 도서관 포털에서 제공하기를 기대하는 기능으로 자원탐색 기능, 통합 검색 인터페이스 기능, 분산검색 기능, 콘텐츠에 대한 직접 접근 기능, 이용자 인증 기능, 그리고 개인화 기능을 제시하고 있다.¹²⁾

도서관 포털에 대하여 많은 논의가 진행 중에 있지만,¹³⁾ 통합검색 환경과 관련된 요소기술을 망라적으로 검토하여 통합검색포털 구축을 시도한 사례는 영국의 JISC IE 사례가 가장 좋은 참조 모델이다. 오랜 기간동안 정보환경으로서 도서관 포털에 대한 다양한 요소기술에 대한 검토와 적용이 논의되고 있기 때문이다. 최근의 주요 사례로는 미국의 NSDL의 OCKHAM Library Network가 있다. 이는 NSDL 프로젝트에 의해 구축된 각종 디지털장서들을 전통적인 도서관 서비스에 통합하고자 하는 프로젝트이다. 또한 각종 요소기술(컨포넌트) 기반의 디지털도서관 서비스와 장서들을 생성하는 프레임워크를 구축하여 새로운 디지털도서관 아키텍처를 구상하려고도 한다. 국내의 경우는 광범위한 정보자원을 대상으로 하고, 오랫동안 연구개발되고 서비스되고 있는 국가지식포털 사례가 있다. 이와 같은 사례들을 살펴보면, 현재 거론될 수 있는 통합검색의 유형과 관련된 주요한 방법과 관련된 기법들을 살펴볼 수 있을 것이다.

2. 도서관 포털의 참조사례

가. JISC IE 사례

JISC(Joint Information Systems Committee)는 영국의 고등교육 즉 대학의 강의, 학습, 연구, 행정 등의 영역에서 정보와 통신기술의 효율적인 사용을 위한 각종 프로젝트를 수행하는 조직이

11) Andrew Cox, "Library portals: technical developments," in "LibPortal Project: Access to Library Provided Resources," http://www.lboro.ac.uk/departments/dis/lisu/downloads/LibPortal_final_report.pdf(인용 2006. 4. 25).

12) Ron Davies, "Library portals - today and tomorrow," http://www.online-information.co.uk/2004proceedings/thurspm/davies_r.pdf(인용 2006. 5. 10).

13) 참조할 만한 논의는 다음과 같다. Mary E. Jackson, "The Current State of Portal Applications in ARL Libraries" ARL의 회원도서관을 상대로 도서관 포털의 구현현황을 조사한 보고서이다. www.arl.org/access/portal/PAWGfinalrpt.pdf(인용: 2006. 5. 10). Andrew Cox, Robin Yeates, "Library oriented portals solutions," Chameleon iPortal, DigitalLink/iPac, ENCompass, MetLib, MAP Portal, ZPORTAL 등과 같은 도서관포털 솔루션을 평가한 보고서이다. www.jisc.ac.uk/uploaded_documents/tsw_02-03.pdf(인용 2006. 5. 10).

다.¹⁴⁾ JISC의 목적 중 하나는 망라적인 규모의 학술 및 연구 자원의 장서들에 대한 편리한 접근을 제공하는 영국의 온라인 정보환경을 구축하는 일이었다. JISC의 여러 사업에는 이 목적에 동의하는 미국 등 세계 각국의 관련 조직과 기관이 참여하고 협력하고 있다. 이것이 바로 '정보환경'으로 해석되는 IE(Information Environment)이며, 영국을 배경으로 하는 방대한 규모의 도서관 포털을 연구개발하는 환경이다. JISC IE는 대학이 온라인 자원과 서비스의 가치를 충분히 활용 하는데 필요한 서비스, 도구, 메커니즘 등을 제공한다. 즉, 대학과 관련된 개인이나 기관의 학습, 강의, 연구 활동에서 필요한 온라인 정보자원에 접근하고 활용할 수 있도록 해 주는 것이다.

1990년대 후반부터 JISC는 고등교육 영역에 가치가 있는 다양한 유형의 정보자원과 서비스(서지, 원문, 이미지, 비디오, 지리정보, 데이터세트 등)의 장서에 대한 통합접근 환경을 구축하는데 관심을 가졌으며, 그 결과 DNER(Distributed National Electronic Resource)라는 새로운 정보접근 전략을 수립하였다. DNER의 용어가 뜻하는 바는 'JISC가 지원하는 각종 정보와 관련된 활동과 서비스'¹⁵⁾를 의미한다. 이러한 DNER 개념은 다양한 관련 활동이나 서비스를 개발하는 데 있어 공식적이고 일관된 전략으로 역할을 하게 된다.

DNER를 지탱하는 두 가지 요건은 콘텐츠와 환경이다. 콘텐츠는 높은 품질의 정보자원 장서이며, 환경은 이러한 정보자원 장서에 접근하는 온라인 공간 즉, 정보서비스 환경을 말한다. 2001년부터는 DNER는 통합정보환경이라는 의미가 보다 명확하게 표현되는 IE로 부르게 되었으며, 사람들이 DNER 내에 있는 자원들을 검색하고 접근하고, 사용하고 출판하는 등의 활동을 위한 통합서비스의 개발에 더욱더 큰 관심을 기울이게 된다. 그러므로 학술연구정보의 국가적 수준의 포털로서 역할을 하는 IE를 위해서는 각종 서비스, 도구, 표준, 프로토콜과 같은 기술적인 메커니즘의 제공을 강조하게 된다.¹⁶⁾

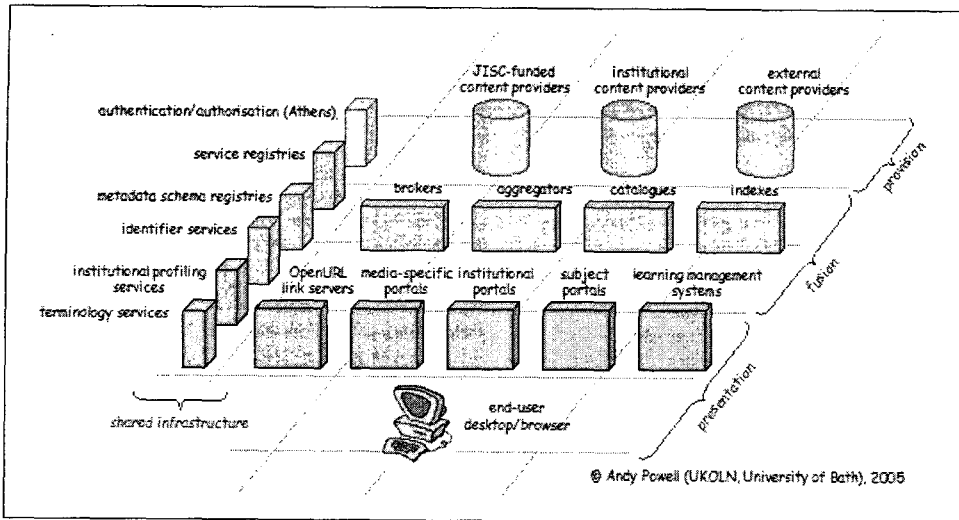
따라서 통합정보검색 환경으로서 IE는 특히 대학 환경에서 학습, 강의, 연구, 행정을 지원하는 질 높은 정보자원에 대하여 안전하고 통합적인 접근을 제공하기 위하여 설계된 것이다. 영국 대학의 이용자들이 출판된 정보자원의 가치를 최대한 활용할 수 있도록 하기 위하여, 현재 존재하는 다양하고 복잡한 유형의 정보시스템과 서비스들로부터 통합한 온라인 정보환경 즉, 정보검색 포털인 셈이다.¹⁷⁾ <그림 2>는 JISC IE의 통합검색과 각종 서비스 제공에 사용되는 표준과 프로토콜들을 설명하는 기술적 아키텍처를 나타내고 있다.

14) <http://www.jisc.ac.uk/>(인용 2006. 4. 15).

15) Stephen Pinfield, Lorcan Dempsey, "The Distributed National Electronic Resource(DNER) and the hybrid library," *Ariadne* 26, 2001. <http://www.ariadne.ac.uk/issue26/dner/intro.html>.(인용 2006. 1. 15).

16) <http://www.jisc.ac.uk/ie/>(인용 2006. 3. 20).

17) <http://www.jisc.ac.uk/dner/development/IEstrategy.html>(인용 2006. 3. 20).



〈그림 2〉 JISC IE 기술적 아키텍처

이러한 아키텍처를 구성하는 주요한 요소기술과 표준화 영역은 다양하다. 이를 영역별로 나누어 정리하면 〈표 2〉와 같다.18)

〈표 2〉 JISC IE의 주요 요소기술과 표준

영역	요소기술 및 표준화
웹 표준 및 파일 포맷 (Web standards and file formats)	HTTP, WAI, HTML/XHTML, CSS, DOM, URI, IMS Content Packaging Specification, METS
분산검색(Distributed searching)	Z39.50, SRW/SRU, Bath Profile, Dublin Core, LOM
메타데이터 수확(Metadata Harvesting)	OAI-PMH, OAI-PMH DC XML schema, LOM
뉴스와 공지(News and alerting)	RSS
문맥기반의 참조연계(Context-sensitive linking)	OpenURL
처리형 서비스(Transactional services)	SOAP, HTTP, XML
이용자 인증 및 접근권한관리 (Authentication and authorisation)	Athens access management system
메타데이터 사용 지침(Metadata usage guidelines)	Dublin Core
서비스 등록기(Service registry)	JISC IESR

이 표에서 알 수 있듯이 IE 아키텍처는 9가지 영역에서 관련된 기술 표준으로 구성된다. 이러한 기술적 아키텍처를 기반으로 하여 IE를 통해 이루어지는 다양한 활동은 다음과 같다. 1) 상호대차, 2) 다양한 검색 서비스(RDN 주제 포털, 기관 및 사업 포털, 개인참조 관리기)를 통한 하위, 원거리 정보자원들의 통합, 3) 탐색 서비스에 이은 제공 서비스의 적절한 연결(link), 4) 정보소스, 학습객체보존소(learning object repositories)의 가상학습 환경과의 통합, 5) e-프린트와 기타 시스

18) <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/standards/> (인용 2006. 3. 27).

템으로의 개방접근(Open Access).

한편 <그림 2>의 JISC IE의 요소기술로서, IESR로 부르는 서비스등록기(services registries)는 JISC IE에서 사용가능한 전자자원(장서)에 대한 목록정보를 등록하고 관리하는 시스템이다. 여기에는 전자자원 장서 그 자체에 대한 정보, 그것에 접근하는 방법에 대한 정보, 그리고 장서의 제공자에 대한 정보 등이 포함된다.¹⁹⁾ 즉, JISC IE에서 제공 가능한 디지털장서를 '장서(collection)-서비스(service)-에이전트(agent)' 엔터티들로 구분되는 각종 지식정보를 등록관리한다. 이것은 RSLP CD 스키마를 이용한 장서 메타데이터에 의해 표현된다.

장서 엔터티는 장서의 명칭, 식별자, 설명문, 유형, 접근권한 등과 같이 장서 그 자체에 대한 정보로 기술된다. 장서는 하나 이상의 서비스를 제공하는데, 이 서비스에 대한 정보는 주로 접근(OAI-PMH, Z39.50, API 등)이나 공지(alert)와 같은 서비스 엔터티에 의해 기술된다. 즉, 서비스 기술 메타데이터는 서비스의 명칭, 식별자, 접근방법, 관리자 등의 정보로 기술된다. OpenURL 해석기와 같이 장서를 갖추지 않은 서비스도 가능하다. 한편, 에이전트 엔터티는 장서를 소유하거나 서비스를 운영하는 기관이나 단체에 대한 정보를 기술하는데 사용된다. 이처럼, JISC IESR는 OAI-PMH, OpenURL, RSS, SOAP, SRU, SRW, Z39.50, Web page 등과 같은 다양한 유형의 프로토콜이나 요소기술을 사용하여 통합검색의 서비스를 제공하고 있다.²⁰⁾

나. NSDL의 OCKHAM 프로젝트 사례

NSDL은 2000년부터 미국 NSF에 의해 구축되었으며, 누구든지 무료로 접근 가능한 디지털도서관이다. 과학, 기술, 공학, 수학(STEM) 분야의 고품질 정보자원과 도구들을 구축하여 강의, 학습, 연구 목적의 접근을 지원하고 있다. STEM 콘텐츠는 다양한 유형의 디지털도서관, NSF 지원 프로젝트, NSDL 인증 웹사이트 등으로부터 수집된 통합장서들로 구성된다.²¹⁾

NSDL은 NSF의 지원으로 현재 약 200개 이상의 다양한 프로젝트들을 수행하였거나 수행 중에 있다. 이들 프로젝트들은 다음의 4가지 영역의 트랙으로 구분된다. 콘텐츠와 장서(content and collections) 트랙은 STEM 분야의 고품질 정보자원 장서를 구축하는 프로젝트로 대부분의 NSDL 프로젝트가 여기에 해당된다. 서비스(services) 트랙은 구축된 콘텐츠와 장서와 관련된 각종 디지털 서비스들을 개발한다. 중점연구(targeted research) 트랙은 서비스나 콘텐츠에 대한 실험작업, 디지털도서관의 분석과 평가와 같이 특정 목적의 연구작업을 수행한다. 핵심통합(core integration) 트랙은 이용자가 콘텐츠, 장서, 서비스에 접근할 때 요구되는 상호운용성을 위한 통합 프로젝트를 담당한다.

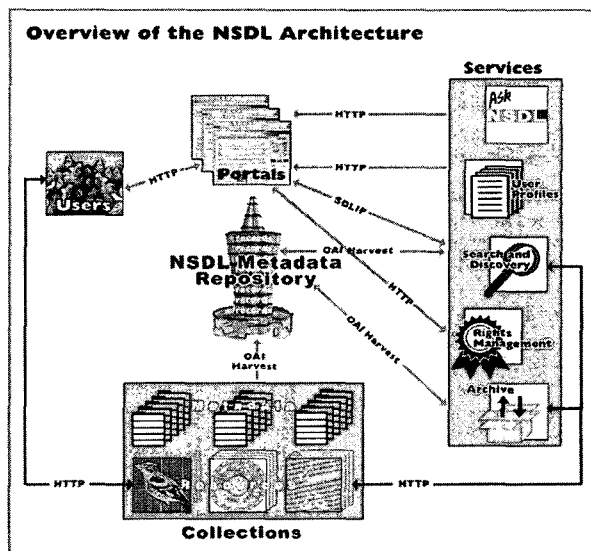
19) Amanda Hill, "The Information Environment Service Registry: Promoting the use of electronic resources," *Ariadne* 40, 2004. <http://www.ariadne.ac.uk/issue40/hill/> (인용 2006. 05. 12).

20) 표준 프로토콜과 요소기술은 IESR의 메타데이터에서 제공되는 서비스의 접근방식(access method)을 참조한 것이다. <http://www.mimas.ac.uk/iesr/metadata> (인용 2006. 5. 10).

21) NSDL의 공식 사이트는 <http://nsdl.org>이다.

OCKHAM²²⁾ 프로젝트는 NSDL의 서비스 트랙에 해당되는 프로젝트로서, NSDL 포털을 전통적인 도서관 서비스에 통합하려는 의도에서 2003년부터 2006년까지 진행되고 있다. 컴포넌트 기반에서 디지털도서관의 서비스와 장서를 생성하는데 유용한 프레임워크를 구축하는 것이다.²³⁾ 즉, 각종 도서관에서 NSDL의 콘텐츠나 서비스에 접근하고 활용할 수 있도록 하는 환경을 제공한다. 이를 위해 디지털도서관 구축을 위한 일반화된 모범사례나 지침들로 구성된 참조모델을 제공하고, 이 참조모델을 활용하여 NSDL의 자원과 서비스를 이용하여 전통적인 도서관의 정보역량을 강화시키는데 필요한 각종 테스트베드 서비스²⁴⁾를 제공한다.²⁵⁾

이러한 프로젝트를 통해 NSDL 포털 아키텍처는 <그림 3>과 같은 개념도로 구성되어 있다. NSDL의 각종 장서들이나 아카이브의 메타데이터를 OAI-PMH를 통해 수집하여 통합 메타데이터 DB를 구축하게 된다. 이용자는 HTTP나 SDLP와 같은 간편한 프로토콜로 구성되는 NSDL 포털을 통해 참고서비스, 검색 및 탐색, 공지정보 등의 서비스에 접근하며, 각종 장서의 콘텐츠에도 접근이 가능하다.



<그림 3> NSDL 포털 아키텍처²⁶⁾

22) 프로젝트의 명칭은 Open Community Knowledge Hypermedia Administration and Metadata의 약칭이다.
 23) 이를 위한 OCKHAM 개발의 기본원칙은 다음 4가지로 제시하고 있다: 개방형 표준 사용(Open Standards), 컴포넌트 기반의 개발(Component Based Development), 간편한 적용(Ease of Adoption), 참여와 협업(Open Collaboration).
 24) 테스트베드 서비스 유형에는 다음의 7가지를 제시하고 있다: 탐색과 접근 서비스, OAI/Z39.50 검색 서비스, 공지 서비스, 브라우징 서비스, 변환 서비스, 목록작성 서비스, NSDL 탐색도구(pathfinding) 서비스.
 25) Jeremy Frumkin, "The problem of mainstreaming digital libraries," *OCLC Systems & Services: International Digital Library Perspectives* 20(3), 2004: 106-109.
 26) <http://comm.nsd.org/download.php/326/NSDLArchOverview.doc> (인용 2006. 5. 10).

다. 국가지식포털

국가지식포털(www.knowledge.go.kr)은 우리나라의 국가기관들이 보유하고 있는 자료(국가지식정보자원)들을 온라인을 통해 공유할 수 있도록 전산화한 국가지식자원에 대한 통합검색 환경을 제공해 주는 시스템이다. 이전에는 국가지식정보통합검색시스템(줄여서 국통검)이라는 다소 긴 명칭을 사용하였지만, 최근 들어 국가지식포털(Korea Knowledge Portal)이라고 명칭을 바꾸었다. <그림 4>는 통합검색, 전문영역별 개별검색, 그리고 디렉토리 검색의 기능으로 서비스되고 있는 국가지식포털의 홈페이지 화면이다.



<그림 4> 국가지식포털 홈페이지

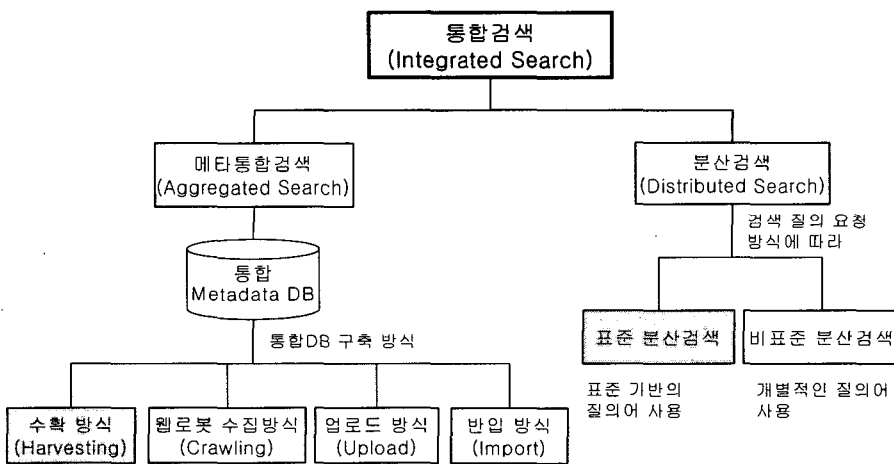
2000년에 제정한 <지식정보자원관리법>에 따라 정보통신부는 국가적으로 보존·이용 가치가 있는 학술, 문화, 과학, 기술 자료를 전문기관으로 나누어 국가지식정보자원으로 지정하고 DB를 구축하여 관리하기 시작하였다. 이렇게 구축된 전문기관의 국가지식 데이터베이스들은 국가지식포털에 연계되어 언제 어디서나 편리하게 검색할 수 있게 된다. 한국전산원을 거쳐 현재 국가지식포털의 전담기관은 한국정보문화진흥원(KADO)에서 담당하고 있다.

전문기관의 통합검색과 관련하여 주목할 만한 사실은 10여개 전문기관이 소유한 약 1,000여개 데이터베이스의 메타데이터를 수집하여 통합한 통합메타데이터 DB를 사용한다는 것이다. 이전에는 메타검색 방식을 사용하였지만, 메타데이터 통합검색 방식을 사용한 이후 검색의 효율성이 매우 높아졌다고 한다.²⁷⁾

국가지식포털에서 제공하는 통합검색서비스는 통합검색, 필드검색, 상세검색서비스, 사이트검색 결과를 제공한다. 통합검색 서비스는 연계기관의 메타데이터를 수집하여 통합메타데이터를 구축하여 이를 기반으로 서비스하는 통합메타데이터 검색서비스이다. 분야별 검색 서비스는 연계기관별로 검색결과를 제한하여 서비스하는 검색서비스이다. 상세검색 서비스는 검색범위, 부가검색, 자료범위, 발행일 등의 다양한 검색 조건을 지정하여 검색할 수 있는 검색서비스이다. 마지막으로 사이트 검색 서비스는 각 연계기관에 원문정보를 제공하는 전문정보센터 및 관련 기관들에 대한 메타데이터를 구축하여 검색 조건에 일치하는 정보를 제공하는 사이트를 소개하는 서비스를 말한다.

3. 통합검색 방식의 유형 분석

지금까지 살펴본 내용을 토대로 디지털도서관 환경에서의 통합검색에 대한 개념을 요약하면 다음과 같다. 기본적으로 통합검색은 다양한 유형의 장서를 대상으로 하며, 도서관 포털이라는 통합 공간에서 수행된다. 이러한 기본적인 원칙을 배경으로 통합검색은 첫째, 내외부의 장서로부터 메타데이터를 로컬 DB에 통합하여 검색하는 메타통합검색(aggregated search) 방식에 의한 통합검색이 가능하다. 둘째, 메타데이터를 통합하지 않고 분산검색(distributed search) 방식도 가능하다. 이들 각각은 다양한 구현기법에 따라 <그림 5>와 같이 여러 가지 유형의 방식으로 세분화될 수 있다. 각각의 세분화된 방식은 앞서 살펴본 사례에서 제시된 표준이나 요소기술을 참조하였으며, 관련이 있는 구현기법들도 포함하여 정리한 것이다.



<그림 5> 통합검색 방식의 유형

27) ZDNet Korea 뉴스기사(2006-02-24).

메타통합검색 방식은 통합검색 대상의 장서(데이터베이스)로부터 수집된 개별자원의 메타데이터를 로컬 DB로 통합하여 검색하는 경우를 말한다. 즉, 디지털도서관의 로컬 DB에 대상 장서의 개별자원들에 대한 메타데이터를 수집하여 통합 DB를 구성하고, 색인작업을 한 후 검색요구를 처리한다. 그러므로 하나의 디지털도서관 내에 집중된 DB를 검색하므로 검색과정 자체는 매우 단순하지만, 다양한 포맷의 메타데이터들을 통합하는 작업이 다양하며 복잡할 수 있다. 통합 DB를 구축하기 위하여 대상 장서로부터 메타데이터를 수집하는 기법에 따라 다른 방식의 유형으로 세분할 수 있다. 가장 많이 거론되는 경우는 OAI-PMH라는 메타데이터 수확 프로토콜 기법을 사용하는 방식이다.

그러나 이와 같은 표준 프로토콜을 사용하지 않고 메타데이터를 통합할 수도 있다. 예를 들어 웹로봇이 대상 장서나 시스템을 주기적으로 방문하여 메타데이터를 수집할 수 있으며, 디지털장서의 메타데이터를 특정한 기법을 사용하여 업로드하거나, 대상 시스템의 메타데이터를 자동(또는 반자동)으로 반입하는 방식도 가능하다. 웹로봇은 특정한 S/W를 이용하여 메타데이터를 수집하는 방법을 대표한다. 업로드 방식은 OCLC, KERIS의 종합목록(union catalog)과 같이 서지 메타데이터를 수집하는 경우에 많이 사용하는 방식이다. 그리고 반입방식은 메타데이터를 통합하는 기타의 방식으로서, 자동 또는 수동으로 메타데이터를 디지털도서관의 로컬 DB에 반입하는 기법을 말한다.

메타통합검색 방식은 동작이 단순하며, 검색행위가 하나의 디지털도서관에 집중된다는 장점이 있다. 그러나 서로 다른 장서들로부터 각각 다른 유형의 메타데이터 포맷을 단일의 통합용 메타데이터 포맷으로 변환하는 작업이 어려울 수 있다. 메타데이터 요소나 데이터 값의 표기방법이나 코드체계가 달라 표준화하고 정규화하는 작업이 소요된다는 것이다. 그리고 데이터의 최신성을 유지하기 위하여 메타데이터의 수집작업이 빈번하게 이루어져야 한다.

분산검색 방식은 이용자가 요청한 질의어를 디지털장서를 소유하고 있는 내외부의 장서 시스템에 복수의 검색질의어를 전달하여 해당 시스템에서 수행된 검색결과를 반환받아 조합하여 이용자에게 제공한다. 즉, 다양한 유형의 통합검색 대상 시스템이 분산되어 있는 상태에서 검색질의를 요청하여 처리하는 방식을 말한다. 분산검색의 대상이 되는 장서의 시스템은 장소, 포맷, 유형, 적용기술 등에 있어 다양하면서 이질적인 특성을 가진다. 이 방식은 교차검색(cross search), 연합검색(federated search), 메타검색(meta search) 등의 용어로도 사용되고 있다. 검색질의어를 개별적으로 검색엔진을 가지고 있는 여러 개의 장서 시스템으로 동시에 전달한다고 해서 병행검색(parallel search), 동시검색(simultaneous search)이나 방송검색(broadcast search)으로 부르는 경우도 있다.

이 방식은 검색 질의어를 전달하고 결과를 수집하는 구체적인 방식에 따라 표준기반 검색과 비표준 방식의 검색과 같이 두 가지 유형으로 구분이 가능하다. 표준기반 검색은 검색 질의어 요청과 결과 회신에 표준적인 프로토콜이나 요소기술을 사용하는 방식이다. 대표적인 표준이나 요소기술

은 Z39.50, SRU, SRW, Web services(SOAP, UDDI, WSDL), Web/CGI 등이 해당된다.²⁸⁾ 한편, 비표준 검색은 표준기반 검색 방식과 달리 대상 시스템의 고유한 검색방식에 따라 질의 요청과 응답 수신 방식이 구현되며, 주로 전용 S/W를 사용한다.²⁹⁾

분산검색의 방식들은 메타데이터를 수집하는 부담이 없으며, 거의 실시간으로 검색이 수행되기 때문에 데이터의 최신성도 잘 유지할 수 있다. 그러나 대상 장서의 시스템에 대한 각종 지식정보³⁰⁾를 유지하고 있어야 하며, 검색질의어를 전달하고, 결과를 수집하여 조합하는 동작이 복잡하며, 대상 시스템들에게 검색의 부담을 줄 수 있다.

이외에도 도서관 포털에서 제공 가능한 통합검색관 관련된 방식으로 참조연계(reference linking) 기법에 의한 연계검색 방식, 디지털장서 수준의 메타데이터(즉, 장서 메타데이터)를 이용하여 장서 수준의 브라우징이나 개별 정보자원에 대한 통합검색을 제공하는 방식 등이 가능하다. 연계검색 방식은 OpenURL 표준을 이용한다.³¹⁾ 장서 메타데이터(collection metadata)를 이용한 통합검색은 최근 들어 영국, 미국 등에서 관심을 가지면도 등장한 개념으로, 디지털도서관의 다양한 장서들에 대한 효율적인 통합관리와 편리한 검색기능을 제공하기 위해 구상된 것이다.³²⁾

IV. 결 론

그 동안 국내의 디지털도서관 시스템에서 통합검색 문제가 크게 부각되지 않은 측면이 있다. 대부분의 디지털도서관 시스템이 패키지 또는 프로젝트에 의해 개발되었으며, 통합검색에 대한 요구가 시스템의 패키지나 개발 프로젝트의 주요 기능으로 포함되었기 때문이다. 그 기능은 디지털도서관을 운영하는 기관이나 패키지를 보급하는 업체에서 충분히 제어가 가능한 그런 수준이었다. 영국

28) 분산검색 방식은 미국의 NISO가 MetaSearch 프로젝트 연구도 참조할 만하다. 이 프로젝트에서는 접근관리(Access management), 장서 및 서비스 메타데이터 기술(Collection & Service Descriptions), 검색(Search/Retrieval)의 영역에서 관련된 요소기술과 표준에 대한 연구를 진행하고 있다.

http://www.niso.org/committees/MS_initiative.html (인용 2006. 5. 10).

29) Tamar Sadeh, "The challenge of metasearching," *New Library World*, 105(1198/1199), 2004: 104-112.

30) 지식정보는 장서 메타데이터를 의미한다. 장서 시스템의 이름, 포괄범위, 언어, 데이터유형 등과 같은 장서 그 자체에 대한 서지 메타데이터와 프로토콜 유형, 목록유형, 레코드 구조 등과 같은 접근에 필요한 정보기술적인 정보로 구성된다.

31) 이 방식에 대해서는 다음 논문을 참조. 이수상, 이길신, "OpenURL을 이용한 정보자원의 연계기술에 관한 연구," *도서관* 58(1), 2003: 107-132. 김성희, "OpenURL을 이용한 전자자원 링크시스템 비교·분석," *정보관리학회지*, 제22권, 제4호(2005), pp.221-234.

32) 현재 거론되고 있는 표준적인 장서 메타데이터는 RSLP CD 스키마, DCMI CD AP가 대표적이다. DCMI CD AP는 NISO CD 스키마라 부르기도 한다. 장서 메타데이터를 이용한 통합검색 기반을 제공하는 주요 사례는 SCONE, GDL, NSDL, OAC, EnrichUK, IMLS DCC, Cornucopia 등이 있으며, 이들은 대개 특정 지역이나 주제 영역의 도서관, 박물관, 기록관 등의 장서를 대상으로 하고 있다.

이나 미국처럼 다양한 표준화와 요소기술의 기반이 부족하였고, 관련된 프로젝트를 장기적 차원에서 대규모로 시도해 보지도 못한 형편이다.

사정이 어떠한 간에 디지털도서관의 통합검색은 그 대상이 다양해지고, 특히 기존의 시스템과는 아주 이질적인 정보환경을 갖춘 시스템이나 데이터베이스 장서들을 대상으로 하여야 한다. 그러므로 통합검색은 보다 체계적이고 표준적인 통합검색 방법을 사용하여 접근하며, 동일한 통합대상이라고 하더라도 기관이나 이용자의 성향에 따라 다른 방식의 통합검색이 필요하기도 한다.

이 연구는 통합검색에 대한 요구가 다양하다는 전제하에, 적용 가능한 통합검색 방식을 유형별로 구분하고 관련된 표준화와 요소기술의 영역을 살펴보았다. 디지털도서관의 통합검색 방식은 도서관 포털이라는 웹 환경을 배경으로 한다. 그리고 메타통합검색과 분산검색으로 양분하여 접근하고 있다. 전자는 OAI-PMH의 메타데이터 수확 프로토콜을 이용하며, 후자는 Z39.50, SRU/SRW와 같은 표준 프로토콜이나 요소기술을 이용한다. 특히 분산검색 방식은 JISC IE, NSDL OCKHAM, NISO MetaSearch 프로젝트와 같이 표준화와 요소기술에 대한 연구와 시범적용 작업이 진행 중에 있다. 또한 MetaLib, ZPORTAL 등과 같은 상용의 도서관 포털 솔루션에서도 관련된 기능이 구현되고 있다.

한편, 디지털도서관의 통합검색은 현재 다음의 특징을 가지는 방향으로 발전하고 있다. 첫째, 도서관 포털이 연구/강의/학습 영역의 전문적인 정보를 제공하는 전문포털로서의 역할을 갖추어 가면서, 온라인 강의시스템, 인터넷 검색포털 등과 같은 전문정보를 필요로 하는 다른 응용 시스템과 연계하는 방향으로 진화하고 있다. 둘째, 웹서비스 기술에 기반한 표준과 요소기술의 적용, 웹2.0, 시멘틱웹 등을 배경으로 하는 새로운 정보플랫폼이 구축되고 있다.

전자는 정보검색 영역이 상호간 연계와 통합을 지향하는 현상과 관련이 있다. 이 현상의 선두주자는 단연 구글 Scholar(<http://scholar.google.com>)이다. 세계적인 인터넷 검색포털인 구글을 통해 도서관에서 제공하던 각종 전문정보를 검색할 수 있도록 한 것이다. 인터넷 검색포털인 구글과 도서관 포털이 통합검색을 위해 연계한 모델이다. 국내에서는 엠파스의 과학기술검색(<http://science.empas.com>), 네이버 전문지식 검색(<http://km.naver.com>) 등의 활동이 이 사례에 해당된다고 하겠다. 현재 국가 지식포털, 국립중앙도서관, 한국과학기술정보연구원, 국회도서관 등의 정보자원은 이러한 포털들을 통해 접근할 수 있도록 하고 있다. 이처럼 도서관 포털은 자체적인 서비스뿐만 아니라 외부의 서비스 응용에서 쉽게 연계되고 활용될 수 있도록 개방화 되고 있다. 도서관 포털들이 서로 연계하여 하나의 거대한 정보 네트워크가 구성되어 가고 있는 것이다.

전자의 특징이 도서관 포털의 검색 방식에 대한 변화라면, 후자의 특징은 이러한 변화를 주도하는 기술이나 인간과 관련되어 있다. 즉, 웹 자체가 사이트나 포털의 공간을 제공하던 역할을 넘어 정보 플랫폼을 구성하고 있다. 이것은 차세대 웹으로 각광받고 있으며, 인간 중심의 참여, 공유, 개

방이라는 키워드를 강조하는 웹 2.0(Web 2.0)과 관련이 있다.³³⁾ 주로 SOAP, SRU, XML 등 웹 서비스 요소기술과 OAI-PMH, OpenURL 등과 같은 디지털도서관의 핵심 프로토콜을 배경으로 하는 것으로 보아, 웹 2.0과 디지털도서관은 통합검색을 위한 플랫폼을 지향하고 있음을 알 수 있다. 디지털도서관은 이용자의 참여, 정보자원의 공유, 그리고 개방된 통합검색 플랫폼으로 구축되어 질 것이다.

이처럼 디지털도서관의 통합검색은 지속적으로 변화하고 발전하고 있다. 그에 따라 디지털도서관의 서비스 방식에도 많은 변화가 예상된다. 통합검색 대상의 디지털장서의 유형과 포맷이 다양해지는 것은 물론이고, 박물관이나 기록관, 그리고 공공영역이나 상업적 영역의 전문정보를 통합하고 연계하는 거대한 플랫폼인 정보환경을 구축하게 될 것이기 때문이다. 기관 중심의 통합검색, 지역 기반의 통합검색, 그리고 특정한 주제 영역의 통합검색과 같이 분화할 수도 있다. 도서관계는 이와 같은 다양한 변화에 대하여 신중하게 대응해야 하지만, 관련된 표준화와 요소기술의 연구개발에 소홀해서는 안될 것이다.

〈참고문헌은 각주로 대신함〉

33) 웹 2.0은 시멘틱웹과 비슷한 개념이다. 웹을 바라보는 관점에서 차이가 있다. 웹2.0은 웹을 인간 중심에서 바라보며, 시멘틱웹은 웹을 기술 중심에서 바라본다. 즉, 웹2.0에서 웹은 새로운 기술보다 인간이 웹을 사용하고 바라보는 관점에서 사용하는 개념이다. 도서관 영역에서 웹 2.0은 이제 막 논의가 시작되고 있으며, 사람들에 따라서 '도서관 2.0(Library 2.0)'이라는 명칭을 사용하기도 한다. 이와 관련된 최근의 논문은 다음과 같다. Theo van Veen, "Serving Services in Web 2.0," *Ariadne* 47, 2006. <http://www.ariadne.ac.uk/issue47/vanveen/>. Paul Miller, "Coming Together around Library 2.0," *D-lib Magazine*, Vol.12, No.4(2006). <http://www.dlib.org/dlib/april06/miller/04miller.html>. (인용 2006. 5. 12).