

가상 카탈로그를 통한 정보공동체의 실현

정보의 축적 및 검색 등 위해 글로벌 작업 필요할 때

어니스트 페레즈(Ernest Perez)

전문 도서관원들은 오랫동안 “가상 컬렉션”을 만드는 작업을 진행해왔다. 우리는 전통적으로, 실제 항목들의 대규모 합동 컬렉션들을 인위적으로 구축함으로써가 아니라 소장자들의 도서목록 기록을 축적시킴으로써 도서관 자료들을 결합시키는 작업에 접근해왔다. 인터넷 이전의 시대에는 이를 가상 카탈로그라기보다는 “결합” 카탈로그라 칭했었다.

도서관원들은 또한 DDC, LCSH, AACR, 시어스(Sears)등과 같은 표준 및 프로토콜들을 매우 성실히 준수해왔다. 이의 목적은 일관성 외에도, 언제나 표준 카탈로그 데이터셋 및 기록의 신뢰할만한 공유였다. 컴퓨터 시대에 접어들면서 이는 원격 데이터 접속 및 데이터 기록의 전달을 의미하는 것으로 확대되어 왔다. 오늘날 우리는 ACSII, Z39.50, ODBC 및 기타 컴퓨터화된 다양한 포맷이나 프로토콜들에 노력을 집중하고 있다.

표준에 집착하는 이유

이는 단지 데이터 전달상의 편의에 관한 문제가 아니다. 물론 데이터 이식가능성에 관한 한 여러분은 여러분의 카탈로그를 블랙홀에 집어넣기를 원하지 않는다. 다른 모든 데이터베이스 어플리케이션들과 마찬가지로 하드웨어 및 소프트웨어 테크놀로지와 보조를 맞추기 위해 영구적인 가치를 지니는 온라인 카탈로그 데이터 파일이 마침내는 보다 새로운 시

스템으로 이전 또는 전달되어야 할 것이라는 것은 피할 수 없는 사실이다.

저자가 이전의 칼럼들에서 지적했듯이 여러분은 세상에서 가장 큰 데이터 파일을 가지고 있을 수도 있으나 여러분의 정보 제품은 그것이 PC-DOS나 C/PM 시스템과 같은 구식 테크놀로지 플랫폼에 보관되어있다면 분명 어려움을 겪을 것이다.

그러므로 지각 있는 기획자라면 일부 비-프로토콜의 장래성 없는 시스템에 시간과 노력과 돈을 허비해서는 안되겠다.

새로운 테크놀로지와 발을 맞추는 것은 분명 다른 시스템들에 있는 정보에 대한 쉽고 투명한 접속의 편의성을 추가해준다.

TCP/IP, Z30.50, ODBC는 모두 그러한 혜택의 예들이다. 그러나 정보의 누적적 가치 또한 이러한 표준들을 사용함으로써 얻을 수 있는 또 다른 중요한 혜택이다.

분포된 정보 자원을 통일시킴으로써 상당한 시너지와 혜택들을 얻을 수 있다. 개별적인 데이터 파일들을 결합시키면 데이터베이스나 카탈로그 어플리케이션들에 옥탄가 높은 동력 지원을 제공할 수 있다. 우리 모두는 OCLC나 WLN과 같은 결합된 도서목록 유틸리티들의 힘을 맛보아왔다. 이러한 시스템들은 상이한 여러 기고기들의 기록들을 단일의 메가-데이터베이스로 결합시킨 연합 데이터베이스로서 구축되었다. 그러나 현대의 네트워킹, 프로세싱, 소프트웨어 기능들은 가상으로 축적되는 개별 또는 분포 데이터베이스의 가능성을 열어주고 있다. 도서관의 경우 이것이 바로 가상 카탈로그인 것이다.

날아다니는 축적

원격 정보 자원에 대한 쉽고 저렴한 인터넷 연계가 어떻게 이러한 혜택들을 제공할 수 있을까? 이는 단순히 신속 원격 접속의 경우가 아니다. 다음 단계로서 원격 컴퓨터 서버에 접속함으로써 다른 원격 컴퓨터들에 접속하여 동시에 상호 작용하는 매개체로서 그 컴퓨터상의 프로그램을 사용하는 방법을 생각해보자.

그런 다른 시스템들에 직접 접속하는 데 약간의 시간이 더 소요되는 것이라면 왜 굳이 이렇게 해야하느냐고 여러분은 반문할 수도 있다. 그건 그렇지가 않다. 여러분 또한 여러분이 직접 할 수 있는 모든 일들을 정말 신속하고 오류 없이 처리할 수 있는 소프트웨어를 사용함으로써 전형적인 상층작용을 얻게 된다. 3개나 5개 또는 n개의 다른 컴퓨터들과 동시에 상호작용하는 것 외에도, 이러한 소프트웨어는 꽤 많은 수의 다른 작업들을 동시에 처리할 수 있다. 예를 들어 한 프로그램은 다른 시스템들에 대한 투명한 검색을 위해 프로토콜-포맷 질문들을 사용할 수 있다. 이는 데이터를 다운로드 하면서 프로그램된 상호작용 시퀀스를 처리하고 진행 중 프로세싱을 행하면서 누적된 결과들을 통일 또는 표준화할 수 있다. 이는 또한 다수의 리소스 검색에서 얻은 결과들을 중복삭제, 분류, 전달하여 최종 결과로서 하나의 편집된 고품질 정보 상품을 전달할 수 있다. 중앙 서버상의 이런 류의 프로그램은 말 그대로 "가상 매개체"이다. 그리고 이는 고도로 숙련되고 능숙한 매개체이다. 만일 여러분이 이것이 공상과 학소설에 가깝다고 생각한다면 수많은 도서관에서 현재 이행되고 있는 가상 카탈로그 방식을 살펴보도록 하라.

이제 더 이상은 공상 속의 제품이 아니다

CPS Systems, Inc.와 OCLC은 모두 비독점 OPAC들을 연결시키기 위한 가상 카탈로그 시스템을 개발중에 있다. 이는 자동화된 지역 ILL 네트워크를 향한 진일보이다. Innovative Interfaces 또한 유사한 시스템을 제공하나 이는 이 회사가 소유한 독점 온라인 카탈로그 설치로 제한되어 있다. 오레곤 주립 도서관에서 우리는 CPS Systems의 URSA(일반 리소스 공유 어플리케이션) 시스템을 경험하였다. 우리는 'ORION' 지역 ILL 검색 및 트랜잭션 시스템용 LSTA-기금 기본형 구성으로 이를 사용하였다.

URSA는 가상 카탈로그 및 자동화된 ILL 트랜잭션 시스템이다. 이는 인터넷상의 다수의 원격 OPAC 호스트들을 동시에 검색하여 누적되고 표준화된 검색결과 세트들과 디스플레이를 신속하고 정확하게 전달할 수 있다. 이는 다양한 OPAC 판매상들과 공조한다. 이는 직원-중재 모드로 운용될 수 있으며 아니면 도서관 고객들이 직접 사용할 수도 있다. 또한 미세조정하여 사용자-개시 리퀘스트를 허용하거나 상이한 여러 수준의 직원 허가를 요할 수 있다.

URSA는 참여 호스트 OPAC와 URSA UNIX 워크스테이션간의 인터넷 전용 접속을 요한다. 이는 모든 또는 선택된 OPAC들의 동시 검색, 대역 도서관 OPAC에서 보유 및 끌어오기 리퀘스트 배열, 요청 도서관에서의 고객 상태 점검, 대역 도서관에서 해당 항목 요금 징수, 임시 항목 기록 생성 및 요청 도서관에서 해당 항목 요금 징수, 리퀘스트 및 항목의 동적 또는 보고서 상태 제공 등을 포함하여 접속된 호스트 OPAC 사이의 실제적으로 모든 ILL 운영을 자동화시킬 수 있다. 그러나 도서관이나 고객 최종 사용자들은 네트워크되거나 전화회선을 이용한 인터넷 접속을 사용하여 검색하기 위해 간단한 웹 브라우저 인터페이스를 사용하기만 하면 된다. URSA는 또한 호스트가 아닌 도서관들을 위한 ILL 차용자 기능을 제공하는데 이런 도서관들은 완벽한 인터페이스를 위해 간단한 웹 브라우저를 사용한다.

URSA같은 시스템들은 이미 언제든 구입 가능한 상태이거나 제품 배달 단계에 있는 상태이다. 이러한 제품들은 여전히 주문제작되거나 1회 한정 시스템들이므로 플러그를 꽂아 작동되거나 완성 인도 방식의 배달 상태에 있지 않음이 분명하다. 가상 카탈로그 실행은 협동 시스템에의 사이트 참여업체들간의 진정한 노력을 필요로 한다.

물론, 첫째는 표준이다. 일찍이 여기 오레곤에서 우리는 주요 영역들에 대한 표준에 동의해야 했다. 우리는 자동 질문들이 어떤 정보 요소들과 세부사항들을 검색, 전달, 디스플레이할 것인지 그리고 시스템이 어떤 방식으로 워크플로우를 처리할 것인지에 대한 규격을 채택했다. 이는 기본적으로 많은 상이한 도서관들의 ILL 도서관원들과 시스템 직원들이 하나의 방에 들어가서(여러차례) 동의된 데이터 요소 세트, 데이터 포맷, 디스플레이 프레젠테이션, 분류 순서, 논리 운용 등을 가지고 다시 나와야 한다는 의미였다. 이들은 또

가상 카탈로그 제품들에 대한 웹 사이트

URSA(CPS 시스템) <http://www.cps-us.com>

SiteSearch(OCLC) <http://www.oclc.org/oclc/menu/site.htm>

GALILEO <http://www.peachnet.edu/galileo/about.html>

MnLINK <http://www.mnlink.org>

OCLC Newsletter <http://www.oclc.org/oclc/new/n236/mn-cic-and->

한 일반적으로 수용되는 정보 및 ILL 트랜잭션 워크플로우를 규정해야 했는데 그럼으로써 이들은 시스템 결과의 의미, 미결정 리퀘스트들의 상태, 시스템 운용 단계들의 의미를 이해함에 있어서 모두 동일한 페이지에 있게 되며 또한 심지어는 필요한 작업들을 하고 있게 되는 것이다.

우리가 얻은 보다 쓰라린 교훈은 프로토콜을 준수하는 것으로 모든 것이 조금이라도 해결되는 것은 아니라는 점이다. 이런 사실에도 불구하고 OPAC 판매상들은 변함 없이 "Z39.50 표준의 준수"를 주장하고 있으나 이러한 순종은 위반행위가 횡행하는 가운데 감사히 받아들여지는 경우가 많다. 일부 변형물들은 아마도 제1 독점 제품들의 서로 다른 접근방식들이나 기능들로 인한 것이나 일부는 임의적으로 선택된 일탈인 것처럼 보인다. (반면에 저자는 보다 관용적으로 이들을 우리 미국의 경쟁 시장에서 유래된 창의적인 변형물, 확장물, 또는 "특징"이라 간주하고자 한다.)

시스템들이 서로를 인식하게 만들기

구성 노력의 특히 상세한 한가지 부분은 각각의 참여 호스트 사이트에서 Z39.50 프로토콜 세트 요소들의 가치 및 특성 맵핑이었다. Z39.50 응답 성능 차이는 다음과 같은 2가지 요인에서 기인한다: 1) 특정 OPAC 판매상이 제공하는 제1 Z39.50 컴플라이언스 및 운용 방법, 2) 특정 도서관 설비에 이루어진 구성 선택물들.

첫 번째 경우에는 독점 OPAC 제품들이 근접 또는 구 검색, 주제 또는 제목 키워드 검색, 또는 저자명 검색과 같은 과제들에 대해 동일한 알고리즘이나 프로세싱 논리를 사용하지 않는다. 예를 들어 어느 한 특정 도서관은 특정 분야의 키워드 인덱싱을 선택하지 않았을 수도 있으며 아니면 수 없이 많은 다른 구성 선택안들 중 하나를 나름대로 국부적으로 변형

하기로 결정했을 수도 있다. 그런 국부 사이트에서는 아무런 문제가 없는데, 그 이유는 OPAC가 그 특정 도서관의 희망대로 수행할 것이기 때문이다.

그러나 이렇게 작게 토막난 모든 변형물들의 누적적 효과는 무엇인가? 만일 여러분이 동일한 필드들을 나와 동일한 방식으로 인덱싱하고 있지 않거나 여러분의 OPAC가 특정 운용에 대해 내 것과 동일한 종류의 결과들을 제공하지 않는다면 우리의 정교한 가상 카탈로그 서버는 아웃셋에 비표준 결과들을 가지게 된다. 이는 최종적인 누적 결과들이 완전하지 못하게 되리라는 의미이다.

많은 이들이 소프트웨어가 "가상 연합 카탈로그"를 생산해 주길 희망해왔다. 유감스럽게도 저자는 이것이 가능하리라고 생각지 않는데 그 이유는 부분적으로는 방금 지적했던 OPAC 고유의 변형물들 때문이다. 둘째로는 여러분은 요금을 지불하지 않은 것은 얻지 못할 것이다. 자동화된 연합 카탈로그 운용의 개발에는 분명한 일관성 혜택과 그와 똑같이 분명한 비용이 존재한다. 첫 번째 비용은 분명한데, 참여업체들은 추가 OPAC 설치 및 운용에 대해 요금을 지불하게 된다. 두 번째 비용 항목은 다소 불분명한데, 참여업체들이 또한 도서목록 기록, 주제 표제, 컴퓨터화된 포맷, 그리고 비표준 형식으로 기입한 다른 모든 것들을 표준화하기 위한 노력과 비용을 지원하고 지불하게 된다. 이런 식의 추가로 진행되는 경비는 카탈로그 검색 결과들의 품질 및 일관성으로 확실히 보상받게 될 것이다.

그러므로 ORION 시스템은 보다 정확히 말하자면 평범한 "가상 카탈로그"이다. 이는 또한 ILL 트랜잭션의 효과적인 자동 처리 및 추적을 제공한다. 그러나 진정한 연합 카탈로그 기능을 제공하지는 않을 것이다.

이러한 차이와 비호환성은 구, 키워드, 필드 규격, 날짜 등을 조합해서 하는 복잡한 검색의 경우에 가장 두드러지게 드러날 것이다. 이런 류의 검색은 때로 필요하다. 그러나 문제를 해결할 수 있는 가장 쉬운 방법은 질문을 간단하고 직접적인 것 - 저자, 제목, ISBN/ISSN 등으로 유지하는 것이다. 이는 ILL 및 자원 공유 도구일 뿐임을 명심하라. 이는 총비용의 80%를 들여서 마지막 남은 20%의 품질을 가지는 80/20의 문제이다. 가상 카탈로그는 정보공유의 진일보라 할 수 있지

만, 절대 안전한 검색 및 브라우징, 발견 도구로 진행되지는 않을 것이다.

정말로 얼마나 괜찮은 것인가?

URSA는 환상적인 ILL 트랜잭션 도구이다. URSA 제품의 SuperSearch 설치를 생각해보자. 이 가상 카탈로그는 지난 20개월에 걸쳐 캘리포니아의 노스 베이 협동 도서관 시스템(North Bay Cooperative Library System)에 28개의 도서관들을 연결시켜왔다. 소노마 군 도서관의 시스템 및 기술 서비스 매니저인 짐 로사치(Jim Rosaschi)는 "SuperSearch를 통해 고객이 직접 한 항목을 요청하는 데 우리 직원이 사용하는 시간은 5분이다. OCLC를 사용하여 ILL 직원이 동일한 요청을 하는 데는 27분이 걸렸다"고 전한다. 도서관에서 숙련된 노동력이 1/5로 절감된다는 것은 이 시스템을 채택하는데 대한 영구적인 보상이 될 것이다.

노스 베이 도서관장인 아네트 밀리론(Annette Milliron)은 내게 Super-Search를 사용하여 80-90%의 성공률을 거두어왔으며 이는 계속 향상되고 있다고 말했다. 그는 "성공"이라는 것을 ILL 트랜잭션의 결과 항목이 성공적으로 전달되어 직원이 그 리퀘스트를 전혀 추적하지 않고 처리되는 것이라 정의한다. 물론 이에는 간단한 전달 및 계산 프로세싱은 배제된다. "실패"에는 선반에서 상실되거나 예약된 것으로 반환되어 ILL에 대한 예상으로 이용이 불가능한 항목들이 포함된다. 1998년 그녀가 내게 이처럼 말할 때, 노스 베이는 여전히 최종 시스템 업그레이드를 진행중이었다. 내가 URSA의 성능에 진정 만족하느냐고 묻자, 아네트는 농담삼아 "ILL이라는 옛날 방법으로 돌아갈 길이 없다!"고 대답했다.

협동적으로 ILL를 처리하기 위해 지역 카탈로그를 검색하고 지역 전달 서비스를 사용하는 것 또한 ILL 트랜잭션 비용 삭감을 가져와야 한다. 노동비 외에도 지금 현재 우리는 인접한 지역이나 우리가 거하고 있는 주의 도서관에서 대여하기 위해 OCLC 검색 및 트랜잭션 비용으로 데이터베이스 및 통신 요금을 물어야 한다. 이는 여러분의 지역 TV 방송국 목록을 보기 위해 America Online에 접속해야 하는 것과 같다. 오레곤 주립 도서관원인 짐 셉크(Jim Schepke)는 이를

다음과 같이 표현했다, "우리는 많은 지역 도서관들과 자료를 서로 공유하고 있는데 그 모두를 OCLC를 통해 전달하고 있으며 그에 대해 상당한 금액을 지불하고 있다."

이를 염두에 두면, ILL 트랜잭션 도구들이 ILL 서비스에 서 상당한 연속적인 보상을 가지게 될 것임이 확실하다는 것을 알게된다. 기타 다른 혜택으로 가상 카탈로그는 다음과 같은 것들을 제공한다.

- 도서관 직원들에게 영향을 덜 미치면서 고객의 편의와 속도 증진
- ILL 비용의 경제성
- 많은 노동력 절감 및 ILL 프로세싱에 대한 품질 가치 추가
- 공유 리소스에 대한 크게 향상된 고객 접촉
- 고객 만족 및 대출 부수 증가
- 최종 사용자를 향한 카탈로그 가치 보강, 최종 사용자는 직접적인 혜택과, 지역 리소스들을 매우 신속하게 접할 수 있는 능력을 확인하게 될 것이다.

그러나 너무 느리다

URSA가 완벽하지 않은 것은 분명하다. 놀랍게도 주된 문제들 중 하나는 처음부터 이 모든 것을 가능하게 만들었던 바로 그 인터넷 통신 기능인 것으로 보인다. 인터넷이 또한 최종 응답 시간에 있어서 여러분에게 타격을 가할 수 있는데, 이로 인한 상처는 무시할 수 없는 것이다.

목표 OPAC가 빠르다해도, 그 시스템과의 그 어떤 상호작용에 있어서 인터넷 접속 및 응답에 많은 시간이 소요된다면 효과가 없다. 인터넷 중견들, 지역 WAN, LAN, 라우터, 브리지 이 모든 것들이 지연의 원인이 될 수 있다. 게다가 목표 OPAC가 실제로 응답이 느릴 수도 있다. 이 모든 순간들이 합쳐진다. 그렇게 되면 가상 카탈로그의 "외견상 응답 시간"은 원격 시스템의 실제 응답 시간 중 가장 긴 시간에 인터넷 통신의 전체 양방향 시리즈에 연결되는 시간이 더해진 시간이 될 것이다.

이 둘을 합한 시간은 최종 사용자에게 있어서는 끝없는 기다림으로 느껴질 것이다. 최종 사용자는 현재 자신의 데스크탑 기기상의 번갯불 같은 응답에 익숙해져있음을 기억해야 한다.

분명, 혹자는 n개의 상이한 OPAC들의 동시 질문과 표준화된 누적 결과 세트의 경이로움을 생각하고 이를 높이 평가할 수 있다. 그럼에도 불구하고 23초라는 시간은 길고 또 긴 시간이 아닐 수 없다!

해답은 타협이다. 오레곤의 ORION 기본형 관리자들은 시간에 대해 정확성의 균형을 취하는 의식적인 타협안을 만들었다. 우리는 재빨리 외견상의 시스템 응답 시간이 항상 가장 속도가 느린 시스템의 응답 및 검색 결과 프로세싱에 필요한 시간이 된다는 사실을 깨닫게 되었다. OPAC가 일반적으로 신속한 시스템이라 할지라도, 오늘날의 지루하게 긴 기다림이 인터넷의 보그다운이나 아니면 해당 캠퍼스의 일시적인 라우터 문제라 할지라도, 여전히 최종적인 결과는 참을 수 없을 만큼 느린 응답이다.

우리의 보잘 것 없는 해결책은 최대 대기 시간을 설정하는 것이었다. 이 방법을 사용하면 여러분은 예를 들어 질문을 받은 임의의 OPAC로부터의 최대 응답 대기 시간이 15초가 될 것이라고 결정할 수 있다. OPAC가 그 시간을 초과하면 가상 카탈로그는 그 호스트를 무시해버리고 카탈로그의 나머지에 대한 결과들을 보여준다.

처음에 ORION은 그 카탈로그에 대해 0개의 히트를 디스플레이했다. 그러나 그러한 보고는 도서관이 요청된 항목을 소장하고 있지 않음을 의미하는 것으로 잘못 해석될 수도 있었다. 현재 URSA는 그 시스템을 응답하지 않는 것으로 보고하고 있다. 원한다면 사용자는 해당 도서관 카탈로그를 다음에 다시 검색할 수도 있다.

현실 세계에서의 상황

가상 카탈로그는 아주 잠깐 동안에 전국에서 터져나온 것처럼 보인다. 지금은 개발 단계를 거쳐 상용 유통 단계에 들어가 있다. CPS 시스템들과 OCLC 모두 가상 카탈로그 시스템을 판매하기 위한 마케팅에 힘쓰고 있다. 많은 지역들이 가상 카탈로그 테크놀로지를 실시하고 있는 것처럼 보이나 저자의 생각에는 특정 지역들 외에는 그리 인기를 얻지 못하고 있는 것 같다.

CPS 시스템들은 PALCI, 펜실베이아 대학 도서관 연결 이니셔티브(37개 도서관)의 경우에는 매리랜드주(주전체에 20

개 OPAC)로부터, 오레곤주와 기타 다른 주들의 경우에는 노스 베이 코퍼레이티브(28개 도서관)에서 협동조합들에서 설치 운용되고 있거나 설치되고 있다.

OCLC는 SiteSearch 가상 카탈로그 제품을 널리 설치해왔다. 이 제품은 자동 ILL 기능보다는 가상 카탈로그 측면에 초점을 두고 있다. SiteSearch는 국부 및 원격 온라인 정보 리소스들을 하나의 쉽게 접속하여 이용할 수 있는 전세계적으로 네트워크된 정보 환경으로 결합시킬 수 있는 도구를 제공한다. 이는 회원 도서관들에 다수의 공유 리소스들을 전달하는데 Z39.50 프로토콜을 사용한다. 이들 리소스들에는 FirstSearch와 같은 독점 OCLC 서비스들 뿐 아니라 온라인 카탈로그, 국부 및 원격 데이터베이스, 전문, 영상, 음향 및 비디오가 포함될 수 있다.

SiteSearch는 이들 리소스들에 대한 네트워크 접속을 제공하여 본질적으로 맞춤형태의 전자 도서관을 형성한다. 이들의 WebZ 제품은 ILL 트랜잭션 기능을 제공하며 SiteSearch 프레젠테이션에 포함될 수 있다.

WebZ의 ILL 기능은 ILL 리퀘스트와 트랙킹의 기본적인 것들을 처리한다. 언제든 구입가능한 이는 URSA 제품 고유의 전체적인 ILL 자동화 기능을 제공하지는 않는 것으로 보인다.

이러한 정도의 워크플로우와 프로세싱 자동화가 고객-개시 시스템에 대한 설계 요건이다. 그러나 OCLC는 계약서상에서 특정 컨소시엄이나 집단의 희망에 따라 SiteSearch/WebZ 꾸러미를 맞춤제작할 것이라고 말하고 있다. 이러한 방식은 ILL 자동화와 맞먹는 수준을 이루어낼 수 있다. SiteSearch의 맞춤제작품들은 다음과 같은 사이트들에서 현재 운용중이거나 개발중에 있다.

- 가상 전자 도서관(VEL), 13개 회원을 둔 CIC 리서치 도서관 그룹에 의한 프로젝트

- GALILEO(조지아 대학 및 조지아 주립 대학 도서관들)

가상 카탈로그는 공동체 표준 준수를 보여주고 있으며 정보 솔루션에 실제적인 시너지 효과를 내고 있다. 도서관 시스템 공동체 뿐 아니라 일반적인 컴퓨터 시장이 데이터 포맷 및 상호-플랫폼 통신을 이렇듯 강조하고 있는 것은 매우 다행한 일이다. 