

생명과학 문헌정보 네트워크 프로토타입 설계

안부영, 안성수, 권창혁, 박형선

한국과학기술정보연구원 바이오인포매틱스센터

e-mail : {ahnyoung, ssahn, narrowpath, seonpark}@kisti.re.k

Design of Biology and Bioinformatics Literature Network Prototype

Bu-Young Ahn, Sung-Soo Ahn, Chang-Hyuk

Kwon, Hyung-Seon Park

Center for Computational Biology and Bioinformatic, KISTI

요 약

연구자들이 국외 생명과학관련 문헌정보를 찾으려면 PubMed와 같은 세계적인 문헌정보서비스를 많이 이용하며, 국내 생명과학관련 문헌정보를 찾으려면 KISTI 학회마을, KoreaMed 등 여러 사이트를 방문해야 한다. 이에 생명과학관련 연구를 원활히 수행할 수 있도록 생명정보 데이터베이스와 분석도구를 서비스하고 있는 KISTI 바이오인포매틱스센터(CCB) 홈페이지에서 인터넷상의 Open Access 문헌정보와 국내 학회정보를 수집하여 메타 데이터베이스를 구축하여 서비스하고자 한다. 또한 생명과학관련 주제별 Open Archiving 커뮤니티의 구성과 운영을 통한 연구자간의 정보교환을 유도하고, 더불어 논문뿐만 아니라 세미나, 연구노트 등의 최신의 연구정보를 공유할 수 있도록 본 프로토타입 시스템을 설계하였다.

1. 서론

전세계적으로 유전체연구에 대한 관심과 투자 증대로 유전체 정보의 분석 및 처리에 대한 수요가 급증하고 있는 가운데, 특히 인간의 유전정보를 밝히는 인간 유전체사업(Human Genome Project)이 성공적으로 종결되어 그에 관련된 연구가 매우 활발한 실정이다. 우리나라는 인간 유전체사업에 참여한 다른 나라들보다 생명과학 분야의 연구가 뒤떨어져 있지만 관련 분야에서 두각을 나타내기 위하여 연구에 매진하고 있다.

연구자들이 연구를 하기 위해서는 생명과학 데이터베이스, 대량의 데이터를 분석할 수 있는 분석도구, 그 동안의 연구 결과물이 집적되어 있는 문헌정보가 필요하다. 한국과학기술정보연구원(KISTI) 바이오인포매틱스센터(CCB)에서는 연구자들이 필요로 하는 생명과학 데이터베이스와 분석도구를 홈페이지(<http://www.ccb.re.kr>)를 통해 서비스하고 있

으나, 생명과학관련 문헌정보는 서비스하고 있지 않다.

이에, 바이오인포매틱스센터 홈페이지를 이용하는 이용자들이 생명과학관련 문헌정보서비스를 요청함에 따라, 국내외 생명과학 문헌을 수집하여 메타 데이터를 서비스할 수 있는 문헌정보 네트워크 시스템을 설계하고 프로토타입을 구축하여 이용자들에게 검증을 받고자 한다.

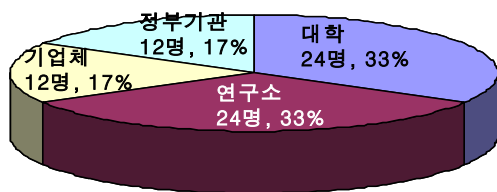
2. 이용자 요구조사 및 결과

본 시스템을 설계하기 전에 한시적으로 바이오인포매틱스센터(CCB) 이용자를 대상으로 요구분석을 위한 간단한 설문조사를 실시하였다. 홈페이지에 팝업창을 띄워서 10일간 설문조사를 실시한 결과 72명이 본 설문에 응답해 주었다. 전체 설문 응답자의 근무기관별 특성은 연구소와 대학이 33%, 기업체와 정부기관은 각각 17%로 나타났다. 전공별 특성은

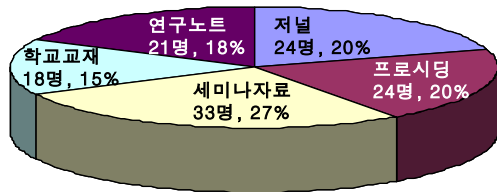
공학을 전공한 응답자가 45%로 가장 많았고 생명과학이 38%, 자연과학이 17%인 것으로 나타났다.

응답자들에게 최근에 이슈로 등장하여 많은 연구가 이루어지고 있는 문헌정보서비스의 기법인 Open Access와 Archiving의 개념을 들어보았냐는 질문에서는 54%가 들어본 적이 있다는 답변을 하였다. 그렇다면 국내에서도 이런 서비스가 필요하겠냐는 질문에서는 들어본 적이 있는 응답자 전원이 필요하다는 답변을 해 주었다. Open Access & Archiving 서비스를 실시한다면 우선적으로 제공되어야 할 문헌은 어떤 것이냐는 질문에서는 세미나자료(27%), 저널과 프로시딩(20%), 연구노트 또는 실험자료(18%), 학교교재(15%)순으로 답변을 하였다.

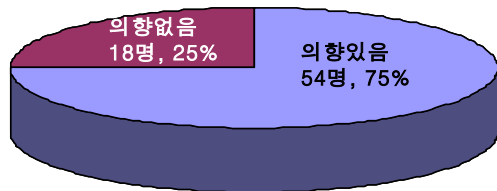
또한 커뮤니티를 구성하여 Open Archiving 커뮤니티를 운영한다면 참여하겠냐는 질문에서는 75%에 해당하는 응답자가 참여하겠다는 의사를 밝혔다. 커뮤니티에 참여한다면 제공할 수 있는 문헌은 어떤 것이 있느냐는 질문에서는 세미나자료가 29%로 가장 많았고, 제공하지 않겠다는 응답자도 29%나 되었다. 이는 제공할 자료는 없지만 필요한 자료는 많다는 결과를 나타낸다고 볼 수 있다.



<그림 1> 근무기관별 특성



<그림 2> 우선제공 문헌자료



<그림 3> 커뮤니티 참여의향

Open Archiving Group에서 공유되는 문헌의 수준은 어느 정도이었으면 좋겠냐는 질문에서는 어느 수준의 자료라도 좋다는 응답자가 53%로 압도적으로 많았고, 국내외 저널정도의 수준이었으면 좋겠다

는 응답자는 21%이었고, 학술발표 자료 수준이면 좋겠다는 응답자는 13%, 세미나 발표자료 수준정도이면 된다는 응답자도 13%이었다. SCI급 자료 수준이었으면 좋겠다는 응답자는 한명도 없었다. 가능한 모든 수준의 자료를 원한다는 결과라고 볼 수 있다.

이상과 같이 설문조사 결과를 알아보았는데, 단기간에 실시한 간단한 설문조사이었지만 본 시스템을 설계하는데 있어 매우 유용한 자료로 활용되었다. 본 설문조사 결과를 바탕으로 국내외에서 공개되어 서비스되는 저널과 아티클을 중심으로 연구자들의 세미나자료, 연구노트, 실험 데이터까지도 공유하고 유통될 수 있도록 본 시스템을 설계하였다.

3. 국내외 현황

3.1. KISTI 과학기술학회마을

KISTI 과학기술학회마을은 국내 과학기술 관련 학회에서 발간하고 있는 학술정보를 데이터베이스 구축 후 인터넷을 통해 서비스하고 있는 학술정보 포털 사이트이다. 2004년 10월 현재 239개의 학회가 등록되어 있고, '학회정보화 지원사업'을 통해 학회 정보화 지원금 지급, 홈페이지 구축, E-Mail 서비스, KISTI-ACOMS(KISTI-학회논문투고관리시스템) 보급 등 학회정보화를 지원하고 있으며, 학회에서 발간되는 학회정보자료를 DB로 구축하고 이를 학회 정보통합검색, 전자화 원문서비스 등을 통해 학회 정보자원의 지속적인 공동 활용체제를 구축하고 있다.

입력, 관리 시스템은 기존의 전통적인 서지입력과 스캐닝의 방식에서 벗어나 학술잡지의 생성 시점에서부터 정보를 관리·서비스 할 수 있는 KISTI-ACOMS 시스템을 개발·보급하여 학회업무의 효율성을 높이고, 정보자원의 최신성 및 신속성을 유지하고자 노력하고 있다.

서비스 주제는 이학분야/ 공학분야/ 농수해분야/ 의약학분야/ 복합학분야/ 기타의 주제로 학회를 분류하고 있으며, 각 해당 저널별 아티클 검색이 가능하다. 각 아티클은 제목/초록/키워드/저자로 나누어 검색할 수 있고, 발행년도에 제한을 주어 검색할 수 있도록 하고 있다. 각 학회목록은 검색이 가능하지만 전체 통합 창을 통해 각각의 아티클 검색이 불가능하다는 단점이 있다.

3.2. KoreaMed

KoreaMed에서는 대한의학학술지편집인협회의 평가 과정을 거쳐 일정 수준이상의 등재 학술지들의

정보를 제공하고 있다. 의학, 치의학, 간호학 학술지 등 총 95종의 영문 초록 7만건을 PubMed 형식으로 제공하며 1997년 이후 출간된 저널들을 the Korean Association of Medical Journal Editors(KAMJE)가 서비스하고 있으며 최근 자료가 95% 이상 수록되어 있다. 전문(全文)은 각 학술지에 연결만 하고 있으며, 1997년 이전의 저널은 KoreaMed Retro라는 서비스로 확장하고 있다. 이 시스템은 서지정보 뿐 아니라 의학 저널 정보도 함께 제공하고 있으며, 2004년 10월 현재 102종의 저널을 링크하여 서비스하고 있다.

3.3. PubMed

PubMed는 미국립의학도서관(US Natinal Library of Medicine)의 NCBI(National Center for Biotechnology Information)에 의해 수행된 프로젝트로 MEDLINE을 가공하여 전 세계에 무료 제공하고 있다. MEDLINE은 생화학 및 의학 분야의 자료들을 중심으로 전 세계의 주요 논문들을 검색할 수 있게 해 주는 데이터베이스로서 이외에도 독성학, 영양학, 약물학, 수의학, 정신의학, 의료공학, 병리학, 스포츠 의학 등에 대해서도 다루고 있다.

많은 서지형 데이터베이스들과 마찬가지로 MEDLINE 역시 학술초록지를 데이터베이스화하고 있는데, 이러한 초록지로는 Index Medicus, Index to Dental Literature, International Nursing Index의 3종을 중심으로 하고 있으며, 이를 통해 전 세계 100여 개국에서 출판되는 3,700여종의 정기간행물에 대한 서지정보(논문제목, 잡지명, 저자, 초록, 출판년도 등)를 제공한다. 1966년 이후의 의학잡지 약 4,650종(MEDLINE: 4,600, PMC: 50)을 색인하고 있으며 1,100여 만건의 Citation을 포함하고 있고, 매월 25,000건 정도가 추가되고 있다.

3.4. DOAJ(Directory of Open Access Journal)

Nordic Conference on Scholarly Communication in Lund/Copenhagen 프로젝트로 시작한 DOAJ(Directory of Open Access Journal)는 Lund university에서 서비스하고 있는 Open Access 서비스로 모든 과학 학술자료를 망라하고 있다. 학교 및 정부, 상업, 비영리단체 등으로부터 자료를 제공받고 있으며 과학기술 연구자들을 Target 그룹으로 삼고 정보를 서비스하고 있다. 이 시스템은 Journal level information, Article level information, 각 Article의

전문(full text) 보여주기(access to full text of articles), 타 서비스와의 통합 연계정보 등의 서비스를 제공하고 있으며, 이용자들이 원하는 메타데이터 정보도 얻을 수 있다. DOAJ의 검색은 알파벳순 저널과 주제별(16개 분야) 저널로 검색이 가능하다.

4. 프로토타입 시스템 설계

생물정보학 문헌정보서비스 네트워크는 아래 시스템 구성도에서 보는 바와 같이 크게 둘로 나뉘어진다. 첫째, 웹상에서 연구자들 간의 정보를 교환·공유할 수 있는 웹 기반 Open Archiving 커뮤니티를 구성하는 것이다. 둘째, 최근 전세계적으로 확산되고 있는 전자저널 중에서 저작권없이 무료로 이용가능한 문헌을 수집하여 연구자들에게 제공하는 Open Access 기반 문헌정보 서비스 시스템이다.



<그림 4> 시스템구성도

본 시스템에 수집되어 서비스될 대상자료는 저널 정보, 발행 전후의 아티클, 학술발표자료, 세미나자료, 연구노트, 실험데이터, 웹사이트, 전자자료 등 연구자들이 필요로 하는 모든 정보가 포함될 것이다.

커뮤니티를 구성하고 문헌정보를 서비스하기 위해 데이터베이스 스키마를 구성하고 메타데이터 레지스트리를 구축하며, 이용자들에게 편의를 제공할 수 있는 키워드를 이용한 통합검색이 가능하도록 시스템을 설계하였다. 본 시스템의 분야별 분류체계는 한국과학재단에서 현대 과학기술의 패러다임을 적용하여 “기초과학연구의 분야분류체계연구” 과제 수행 결과물로 나온 과학기술 분류체계를 사용하였다.

다음의 표는 저널과 아티클 검색결과 화면이다. 최소한의 메타정보만을 출력하고, 타이틀을 클릭하면 원문이 있는 사이트 또는 저널의 사이트로 이동할 수 있도록 하였다. Open Archiving 커뮤니티의 회원들이 올려준 공유 자료의 원문은 회원 로그인을

