

## 생명정보인프라 이용에 관한 연구

### A Study on Use of Bioinformatics Infrastructure

안부영, 이상호  
한국과학기술정보연구원

Ahn bu-young, Lee sang-ho  
Korea Institute of Science & Technology Information

#### 요약

한국과학기술정보연구원(KISTI) CCBB(Center for Computational Biology & Bioinformatics) 웹사이트에서는 IT 기반의 생명정보 인프라 구축을 위해 생명정보 콘텐츠(DB, 분석도구) 21종을 구축·서비스하고 있다. 또한, 국내 생명정보 연구개발 지원 및 인프라 조성을 위한 업무를 수행하고 있다. 하지만 CCBB에 대한 인식이 널리 확산되어 있지 않아 CCBB에서 구축하고 개발된 생명정보 콘텐츠가 활발히 이용되지 못하고 있는 실정이다. 따라서 생명과학 분야의 연구자를 대상으로 이용자조사를 실시하였다. 조사내용은 생명정보 데이터베이스 및 분석도구 이용에 관한 내용과 생명과학 연구 학술정보 네트워크(BioInfoNet)에 관한 내용으로 구성하였다. 설문조사를 통해 나타난 결과는 정리·분석한 후 이용자들이 필요로 하는 데이터베이스, 분석도구 등 생명정보 인프라를 제공하는데 활용될 것이다.

#### Abstract

KISTI(Korea Institute of Science & Technology Information) CCBB(Center for Computational Biology and Bioinformatics) Website has been providing bioinformatics contents(databases and analysis tools) services to construct IT-based bioinformatics infrastructure. It also carries out the services to support domestic bioinformatics R&D activity and to construct the infrastructure. But, these services are not actively available because this CCBB website is not wellknown to users. Therefore, we performed the user survey for the researchers who are studying in the field of biological science. The items of this survey consist of subjects on bioinformatics databases, analysis tools and BioInfoNet. After the user survey is rearranged and analyzed, the result of survey will be used to provide the good service to the users who are need of bioinformatics infrastructure such as databases and analysis tools etc.

## I. 서론

한국과학기술정보연구원(KISTI)은 2002년부터 IT기반의 생명정보 인프라 구축 사업을 추진하고 있다. 생명정보는 인간과 가장 밀접한 대용량·양질의 정보이기에 국가적으로 인프라를 구축하고 제공해야 하므로 국가 과학기술정보유통기관인 KISTI에서 수집·구축·서비스하고 있다.

생명정보학(바이오인포매틱스)은 기초생물학, 의학, 응용생물학 분야에 있어서 필수적인 연구수단이고, 생물학, 전산학, 수학, 물리학 등 타 과학영역간의 연계를 기반으로 하는 연구이므로 생명정보학 연구의 성과는 관련 학문과 산업에 직접적으로 기여할 수 있다. 그리고 생명산업이 사회에 주는 경제적, 사회적 영향이 클 것으로 볼 때 향후 생명정보학은 산업적으로도 가장 중요한 핵심적 요소 기술 중의 하나가 될 것이다.

이에, 생명정보학과 아주 밀접한 학문인 생화학·분자생물학 분야의 학생과 연구자들을 대상으로 생명정보 데이터베이스 및 분석도구와 연구학술정보에 대한 인식과 정보교환 및

공유에 관한 의견을 파악하여 CCBB 홈페이지(<http://www.cccb.re.kr>)에 반영하고자 한다. 또한 KISTI에서 서비스중인 IT 기반 생명정보 인프라 이용의 활성화를 유도하고 그로 인해 향후 국가 생명과학 연구를 지원하는 방안 수립에 적용하고자 한다.

## II. 조사방법

KISTI 사실정보팀은 생명과학 관련 최신 연구동향을 알아보고 CCBB에서 서비스중인 생명정보 콘텐츠 홍보를 위하여 국내 분자생물학관련 최대행사인 한국생화학·분자생물학회 국제학술대회에 전시 부스를 설치하여 홍보 및 이용에 관한 조사를 실시하였다.

2007년 5월 28일부터 30일까지 서울 COEX에서 개최된 제 26회 한국생화학·분자생물학회에는 3,000여명의 국내외 연구자와 학생들이 참여를 하였으며 그중 약 500여명이 전시부

스를 방문하였고 이 중에서 320명에게 설문지를 배포하였다. 그중에 164부가 회수되어, 회수율은 51% 정도였다. 회수율은 저조하였지만 응답자 대부분(85%)이 CCBB 홈페이지(<http://www.ccbb.re.kr>) 회원으로 가입하여 139명의 회원 증원 효과를 얻었다.

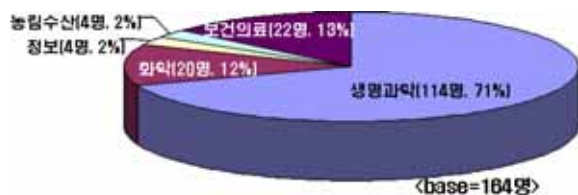
설문 대상자를 본 학회 참가자로 선정한 이유는 생화학이나 분자생물학은 생명정보와는 매우 밀접한 학문이므로 본 학회의 참가자라면 생명정보 관련 최신 연구 및 학문을 많이 접하면서 생명정보에 많은 관심을 가지고 있을 것이라 판단하였기 때문이다.

### III. 조사결과

조사 문항은 크게 응답자 인구통계학적 특성, 생명정보 DB와 분석도구 이용 및 생명과학 연구학술정보에 관한 내용 등 세 부분으로 구성하였다.

#### 1. 조사 응답자 특성

설문 응답자 164명의 주된 특성을 살펴보면 20-30대(90%, 148명) 생명과학을 전공하는 석사과정 이상의 학력을 가진 응답자가 84%(138명)로 가장 많았으며, 57%(93명)가 여성으로 구성되어 있었다. 전공분야는 생명과학이 71%(114명)로 가장 많았고, 화학이 12%(20명), 보건의료가 12%(22명), 농림수산이 2%(4명)이었다. 소속기관은 대학이 65%(107명), 기업체가 18%(29명), 연구소가 11%(18명), 병원은 4%(7명), 정부기관은 2%(3명)인 것으로 나타났다.



▶▶ 그림 1. 응답자 전문분야별 특성

#### 2. DB 및 분석도구 이용

CCBB 홈페이지의 데이터베이스(Genbank, PDB, PIR 등)를 이용하여 보았느냐는 질문에서는 19%(31명)만이 이용경험이 있다고 응답하였으며, 81%(133명)은 이용경험이 없다고 응답하였다. 또한 홈페이지의 생명정보 분석도구(BLAST, FASTA, ClustalW 등)를 이용해 본 경험이 있었는지를 묻는 질문에서는 9%(13명)가 이용 경험이 있다고 답변하였으며, 92%(151명)는 이용경험이 없다고 답변하였다.

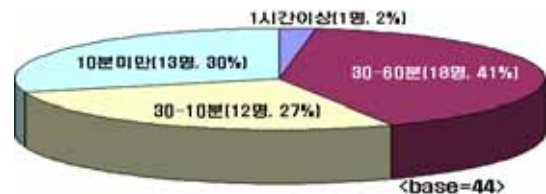
CCBB 홈페이지를 운영한지 수년이 지났고, 국내에서 유일

하게 생명정보 관련 데이터베이스와 분석도구를 서비스함에도 불구하고 대다수의 응답자가 CCBB 콘텐츠(데이터베이스, 분석도구) 이용경험이 없다는 것은 CCBB에 대한 홍보가 많이 미흡하다는 것과 이용자 의견을 반영한 신규 콘텐츠 확충에 더욱 많은 노력을 기울여야 한다는 것을 나타낸다고 볼 수 있다.



▶▶ 그림 2. DB 이용경험

생명정보 데이터베이스와 분석도구의 이용경험이 있는 44명을 대상으로 CCBB 홈페이지를 어느 정도의 주기로 이용하고 있는지를 묻는 질문에서는 하루 한번이상 방문한다는 응답자는 27%(12명), 주 1회 이상은 30%(13명), 월 1회 이상은 36%(16명), 수시로 방문한다는 응답자는 7%(3명)로 나타났다. 한번 방문하면 어느 정도의 시간을 머무는지의 질문에서는 30-60분이 41%(18명)으로 가장 많았으며 30분-10분이 27%(12명), 10분미만은 30%(13명), 1시간 이상은 2%(1명)에 불과했다. 사용자들이 좀 더 오랜동안 수시로 방문하여 홈페이지의 내용을 이용하게 하려면 신규 콘텐츠를 개발하고 최신의 BIT 뉴스 및 동향 등을 전하는 것이 중요하리라고 본다.



▶▶ 그림 3. 홈페이지 1회 이용시간

홈페이지 콘텐츠(데이터베이스, 분석도구)중에서 주로 어떤 것을 이용하느냐의 질문에서는 단백질(41%, 18명)과 유전자(39%, 17명) 데이터베이스를 주로 이용한다고 응답하였다. 분석도구는 9%(4명), 생물자원은 11%(5명)가 이용한다고 응답하였다. CCBB에서 제공중인 생명정보 콘텐츠의 내용이 좋다는 응답자는 43%(19명)이었으며, 32%(14명)은 보통, 20%(9명)은 매우 좋다고 응답하였고, 미흡하다는 평가는 5%(2명)이었다. 응답자 모두가 CCBB 콘텐츠에 만족하는 것은 그동안 지속적인 모니터링을 통하여 콘텐츠를 갱신하고 홈페이지를 이용자 요구에 따라 개선하였기 때문이라고 사료된다. 미흡하다는 평가를 한 응답자 2명에게 미흡한 이유를 물었더니 콘텐

트 부족이라고 답변하였다. 이에, 2007년도에는 이용자들에게 유용한 신규 콘텐츠 발굴과 더불어 서비스를 제공하고자 한다.



▶▶ 그림 4. 주요이용 콘텐츠 현황



▶▶ 그림 5. CCBB 콘텐츠 평가

앞으로 또는 계속 CCBB 홈페이지를 이용하겠느냐는 질문에서는 응답자의 96%(157명)이 이용하겠다고 응답하였으며 4%(7명)은 이용하지 않겠다고 답변을 하였다. 거의 대다수가 이용의사를 밝혔기에 좀더 나은 콘텐츠 제공을 위해 노력해야 할 것으로 사료된다.



▶▶ 그림 6. CCBB 이용 의향

### 3. 연구학술정보 이용

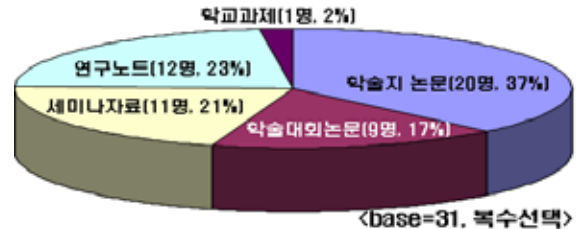
정보의 자유로운 이용이라는 측면에서 새로운 학술정보서비스에 개념을 둔 오픈 액세스 서비스(Open Access Service)에 대하여 들어보았느냐는 질문에서는 19%(31명)가 들어보았다는 응답을 하였으며, 133명(81%)은 들어 본 적이 없다는 답변을 하였다. 이는 생명과학 분야에서는 아직도 저작권과 무관한 학술정보서비스가 활성화되지 않았기 때문인 것으로 생각된다.

그리고 오픈 액세스 서비스를 들어보았다면, 국내 생명과학 분야에서 이러한 학술정보서비스가 필요하다고 생각하는냐는 질문에서는 응답자 100%(31명) 전원이 필요하다는 답변을 하였다. 이는 국내에서도 연구자들을 위한 양질의 문헌을 아무런 제약 없이 이용할 수 있도록 해주었으면 한다는 의사를 표현한 것이라 할 수 있다.



▶▶ 그림 5. 오픈액세스서비스 인지도

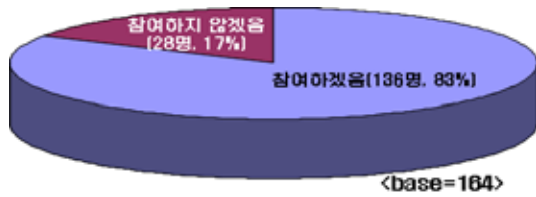
오픈 액세스 서비스가 필요하기에 KISTI에서 이런 서비스를 실시한다면 우선적으로 제공되어야 할 연구학술정보의 종류는 어떤 것이라고 생각하는냐는 질문에서는 응답자 중 37%(20명)가 학술지(journal)에 실리는 논문이라고 응답하였으며, 연구노트나 실험자료는 23%(12명), 세미나 발표자료는 21%(11명), 학술대회(proceedings) 논문은 17%(9명), 그리고 2%(1명)는 학교과제로 제출한 보고서라고 답변하였다.



▶▶ 그림 6. 우선제공 정보종류

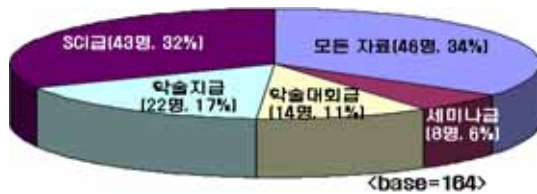
오픈 액세스 서비스가 필요하기에 KISTI에서 이런 서비스를 실시한다면 참여할 의향이 있는냐는 질문에서는 83%(136명)가 참여하겠다는 의사를 밝혔으며, 17%(28명)는 참여하지 않겠다고 응답하였다. 또한 참여한다면 제공가능한 정보는 어떤 것이냐는 질문에서는 33%(62명)가 세미나 발표자료라고 응답하였으며 발행후 6개월이 지난 논문이 28%(51명), 연구노트 및 실험자료가 17%(31명), 학교과제로 제출한 보고서가 14%(25명), 발행전 논문이 8%(15명)로 나타났다. 응답수가 136명이 넘는 이유는 복수응답이 가능한 문항이기 때문이다.

오픈 액세스 서비스에서 공유되는 연구학술정보의 수준은 어느 정도이었으면 좋겠느냐는 질문에서는 어느 수준의 자료라도 좋다는 응답자가 34%(46명)로 압도적으로 많았고, 국내·외 저널정도의 수준이었으면 좋겠다는 응답자는 17%(22명)이었으며, 학술발표 자료 수준이면 좋겠다는 응답자는 11%(14명), 세미나 발표자료 수준정도이면 된다는 응답자는 6%(8명), SCI급 수준이었으면 좋겠다는 응답자는 32%(43명)이었다.



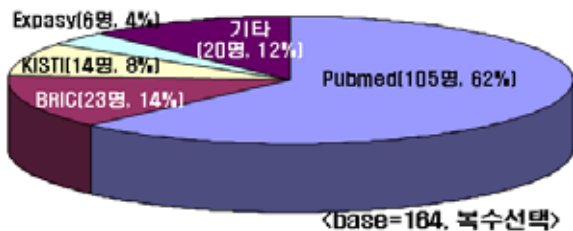
▶▶ 그림 7. 서비스 참여 의향

본 결과를 보면 본인이 제공할 수 있는 정보의 수준과 다른 사람이 제공하는 정보의 수준이 현격히 차이가 나는 것을 볼 수 있다. 이런 요구를 수용하기 위해서는 본 서비스에서 양질의 연구학술정보를 수집하여 제공하는 노력이 필요할 것이다.



▶▶ 그림 8. 공유자료 수준

마지막으로 현재 이용하고 있는 연구학술정보 서비스는 무엇인지를 묻는 질문에서는 62%(105명)가 NCBI에서 서비스하고 있는 Pubmed라고 응답하였다. 국내 서비스로는 포항공대에서 운영하고 있는 생물학연구정보센터(BRIC) 14%(23명), KISTI에서 운영하고 있는 YesKISTI, KOSEN, NDSL이 8%(14명), 바이오인포매틱스 정보가 많은 Expasy가 4%(6명), 학회 및 학교연구실 홈페이지, Scifinder, google, Web of Science 등의 기타 서비스가 12%(20명)로 나타났다. 응답자들이 주로 이용하는 학술정보 서비스가 파악되었기에 위 사이트를 벤치마킹하여 KISTI에서도 양질의 생명과학 연구학술정보를 제공하는 것이 바람직하다고 사료된다.



▶▶ 그림 9. 주요이용 학술정보서비스

#### IV. 결론

이번 설문조사 결과를 바탕으로 사용자들이 주로 사용하는 정보가 무엇인지, 원하는 정보가 무엇이며 어느 정도의 수준인

지, 어떤 서비스를 원하는지를 어느 정도 파악하였으므로 KISTI CCBB에서 생명정보 관련 콘텐츠를 전체적으로 수집하여 서비스할 수 있는 방안을 마련해야 할 것이다. 또한 CCBB를 이용하는 이용자 특성에 맞게 맞춤형서비스를 실시하고 특화된 전문가 집단을 CCBB의 고정 정보 제공자 및 사용자로 확보한다면 양질의 데이터를 수집하고 제공하여 이용자 만족도를 높일 수 있을 것이다.

이에, 이번 설문 조사를 통해 나타난 결과를 정리, 분석하여 기존의 IT 기반 생명정보인프라 구축 사업을 확대하여 사용자들이 필요로 하는 데이터베이스나 분석도구를 신규로 구축·제공하고, 하드웨어와 소프트웨어 인프라를 확충하여 대용량 양질의 정보를 일괄 제공·서비스(FTP 등)한다면 CCBB는 국내 유일의 생명정보 포털사이트 역할을 수행하여 우리나라 생명과학 기술 및 연구개발 발전에 크게 기여하게 될 것이라 기대된다.

#### 참고 문헌

- [1] 안부영외, "CCBB서비스 이용자지침서", KISTI, ISBN 89-5884-066-8 93470, 83p, 2004.
- [2] 안부영, "2006년 생명정보인프라이용 및 만족도조사", 제10회 KOSTI 워크숍 학술발표논문집(III): 427-433, 2006.
- [3] CCBB 홈페이지, <<http://www.cccb.re.kr>>