

# SALIS 2006

*Papers Presented at the National Conference on*  
**Initiatives in Libraries and Information Centers  
in the Digital Era**

*Jointly Organised by*

**SALIS**

**&**

**Karunya Deemed University, Coimbatore**



June 8 - 10, 2006

*Editors*

Dr. Harish Chandra

A. Hariharan

Dr. S. Muralidhar

Dr. J. Dominic

V. Sakthi Regha



Society for the Advancement of Library and Information Science

<http://autolib-india.net/salis>

Chennai

2006

## Contents

Preface

List of Contributors

### Technical Session I - Case Studies

1. Digital Initiatives for Effective Print and Copy Management Solution at BITS, Pilani – Dubai Campus: An integrated approach to access control and Facilities Management 1  
*Kumaresan, S.C. and Sheshadri, K.N.*
2. Implementation of Information Technology (IT) in Library Services with Special Reference to Rajasthan University Library 8  
*Lata Suresh and Prabha Sharma*
3. Enhancing the Usage of Library Resources at SERC Library with Particular Reference to e-Resources 13  
*Hariharan, A., Chitra Hariharan, M., Mymoon, M., Abdunnasar, A. and Suryakala, A.*
4. Role of Library Portal for Online Information Services with Special Reference to Jadavpur University, Centre for Digital Library and Documentation 18  
*Subarna Kumar Das, Sri Rajesh Das, Biswajit Das and Sri Subhendu Kar*
5. Use of Online Public Access Catalogue by the Research Scholars in Annamalai University Library 24  
*David, J.C., Balasubramani, R. and Balaji, N.G.*
6. Training Needs Analysis of the Library Professionals Working in Agricultural Science Universities in India in Emerging ICT Environment: A Sample Survey 30  
*Kannappanavar, B.U., Praveenkumar, Kumbargoudar and Devendra S. Gobbur*
7. Open Source Software: A Case Study 39  
*Vani P Rao and Khaiser Nikham*
8. Use of CD-ROMs and Internet Resources by the Students in JNN College of Engineering, SHIMOGA: a Survey 45  
*Manjunath Lohar., Parameswarappa, V. and Satish K.T.*
9. Facilitating Access and Use of Bioinformatics Information Resources through Digital Library Systems 54  
*Shri Ram and Sanjay Kataria*
10. Evaluative Study of E-resources and Services in Digital Library 63  
*Machendranath, S., Kamble, V.T., Ramesh Gandhi, R.T.D. and Raghavendra, S.*
11. Virtual Library on Poultry Diseases: Design, Development and Evaluation 70  
*Rathinasabapathy, G.*
12. User Behaviour Towards Internet Resources 76  
*Amudha G., Baskaran N., and Lawrence Mary, A.*

# Facilitating Access and Use of Bioinformatics Information Resources through Digital Library Systems

Shri Ram. and Sanjay Kataria

## Abstract

As the Recombinant DNA, Gene Cloning and DNA sequencing technologies improved in 1970s the scientists began to think about the possibilities of the sequencing the  $3 \times 10^9$  nucleotide pairs in human genome. This led to launch of Human Genome Project (1990), the greatest discovery in the field of life science, is near completion. Vast and ever-expanding diverse information on bioinformatics resources is being developed. These include database of biological information, software tools etc. Advances in Internet Technology have largely affected the bioinformatics resources as heterogeneous sources of information. It facilitates the uniform access to the educational, academic and research information sources to the bioinformaticians for their research and developmental activities. This paper describes awareness about the bioinformatics information resources and digital library need of the student at Jaypee University of Information Technology, Solan (HP) and how it can be helpful to access the information resources on bioinformatics through digital library techniques.

**Keywords:** Digital library; Bioinformatics; Web based Information Sources; Information Need.

## Introduction

Over the course of time genetics has become one of the major rapidly growing sciences. The ever-expanding pool of knowledge, molecular biology, databases, non-bibliographic databases information such as DNA Sequence (GeneBank), Protein sequence (Swiss-Prot), and genomic mapping Information (Genome Database) makes it virtually impossible to follow new and recent discoveries using conventional methods such as journal browsing or scientific textbooks. The development, management and use of such databases is one of the major components of the bioinformatics. The computer resources have therefore become critical for management, gathering, and analysis of these information (Missener and Krawetz). Human Genome Project is an international effort initiated in 1990 with the goal to discover all the approximately 30000-35000 genes in human genome and make them accessible for further biological study (National Human Genome Research institute, 2000). Therefore one of the Goals of the Human Genome Project is the development of analysis algorithms and integration of genetic database (informatics) for managing and interpreting genome data.

The Annual Database issue of Nucleic Acid Research Journal listed more than 200 molecular biology databases. The information in these databases include sequences (DNA, RNA and Protein), gene structure, maps, mutation, and bibliographic information among others. The Enterez PubMed, National Library of Medicine USA, provides both Bibliographic and molecular biology database. As the number of and size of molecular databases, bibliographic literature pertaining to these databases grows and their use increases, the associated lacks expertise in the area of bioinformatics to choose right kinds of information at right time. In a study the number of search hits for four defined search queries on some of the more popular search engine and Meta search engines gives out the result as given in Table 1. The table gives the difference between results obtained by the search engine and meta-search engines for particular set of hits. It can be inferred from the above table that the search engines give huge amount of results that student



and scientists community becomes bewildered to choose right kinds of information at right time of requirements. (Lawrence and Giles, 1998).

**Table 1: Result of search hits of search engines and Meta search engines (Lawrence and Giles, 1998)**

	Search Engine					Meta-Search Engine	
Genetic Mapping	478	1040	4326	9395	7043	62	58
Human Genome	13213	34760	15980	19536	19797	42	54
Position Cloning	297	735	1143	666	3987	40	52
Prostrate Cancer	14044	53940	24376	33538	23100	0	57

Faculty members, students, scientists community are overwhelmed by the range and volume of bioinformatics resources available to them, and frequently lacks the expertise to use them (Yarfitz and Ketchell, 2000). There are ranges of research resources and continue to be expanded and developed into complexity and numbers. It is essential to understand the users,

- what information they need from the resources,
- how they search for information
- how they use that information

Earlier research illustrates the information behavior of life science and molecular biology has focused on the use of traditional, bibliographic resources (Bayer and Jahoda, 1981), Curtis, Weller and Hurd, 1993). Some research has addressed the use of non-bibliographic bioinformatics resources. A study made by Yarfitz and Ketchell, 2000 found that 70% molecular biologists surveyed were using molecular biology databases on weekly basis or monthly basis. There are considerable interest in bioinformatics consultation service, classes and other services to provide assistance in the use of bioinformatics resources. Let us define the two terms used above, these are bioinformatics and digital library.

### Bioinformatics

Bioinformatics is a science which developed from integration of biology and Information Technology (IT). The functional aspects of the bioinformatics are the representation, storage and distribution of the data. These goals are realized by intelligent design of data formats; databases populating the databases (Storage of data, creation of tools to query these databases and development of user interfaces about the data) form the basis for scientific collection and storage of data.

The second major aspect of the development of analytical tools is to discover knowledge in data. Biological databases are linkage of complex and interlinked chain of information. The volume of data generated from experiment and stored in databases is increasing exponentially. Till recent, data sources were setup autonomously – websites by individual research laboratories. But major difficulty is that the data elements in various public and private data sources are stored in extremely heterogeneous formats across the internet. These online interfaces for queering and searching the underlying highly nested repository data for research, teaching, industry. The need is to make access the up to date data facilitating faster and meaningful knowledge discovery. To achieve

this goal integrated and fully automated information processing systems such as digital library is required.

### Digital Library

The Digital Library Federation has proposed the following definition:

Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities. This definition emphasizes that a digital library, like any library, is more than a mere aggregation of information resources: it is a service which is based on principles of selection, acquisition, access, management and preservation, related to a specific client community. All of these principles are relevant when we consider the meaning of a digital library and the practical issues involved in service delivery. A digital library collection may include two types of information resource. One type comprises the "digital original" resources, which are sometimes referred to as resources which are "born digitally". The other type comprises "digital surrogates", which are created from traditional information resources through format conversion. While both types of resource have the same access and management requirements, they raise different issues of selection and acquisition, and their preservation imperatives are also different.

Functional aspects of digital library which are required other than standard search and retrieval are necessary for next generations are:

- New operations are required to cope with visualization and manipulation requirement of new type of digital contents. Eg. Genetic Mapping
- When the use of a digital library is closely coupled with users tasks, for example, e-learning, new services that provides seamless integration of digital library access with user tasks will be required.
- Due to advance of web and media technologies, the next generation digital library likely to manage massive amount of digital content. If there are no effective tools to help users locating the relevant information, users are likely to be overwhelmed with too much information, causing the information overload problem.
- To leverage on content and services provided by digital library, interoperability and services among different subsystems of digital library must be supported.

### Agenda for Bioinformatics Digital Library:

The intent of research for bioinformatics digital library is:

- (a) to assess the patterns of use of bioinformatics resources (both molecular biology databases and related analytical software tools) by student, faculty members and research scholars of Jaypee university of Information Technology, Solan, and
- (b) to create predictive model, that will facilitate the access and use of bioinformatics resources by faculty members, research scholar and students. It will help to build a predictive model for the use of bioinformatics resources for a particular scientific task, and use the model to develop a tool to assist scientists in their selection and use of digital information resources available at different nodal centers, institutes and information community.



### Objectives of the Bioinformatics Digital Library

1. To identify and assess the patterns of use of bioinformatics resources by student, faculty members, students of Jaypee University of Information Technology, so that a predictive model can be developed.
2. To create and test a computer-based tool, based on the predictive model, that can assist novices to access and use bioinformatics resources.

### Methodology

To know the access patterns and use of the bioinformatics resources at JUIT, the students were given a set of questionnaire pertaining to the basic information about the bioinformatics. The objective of the questionnaire is to find out the basic knowledge about the bioinformatics information resources, awareness of digital library, their source of information for various curriculum supplements and to assess the patterns of use of bioinformatics resources by student community who are proficient in their use. Six hundred students of the JUIT from the branch of Electronics and Communication Engineering, Computer Science, and Bioinformatics were asked to fill a semi structured questionnaire. Out of those 600 students 135 students of bioinformatics department were subjected to data analysis by use of simple statistics methods of percentage analysis. The following results can be viewed based on the response.

### Awareness of bioinformatics

Table 2: Awareness of Bioinformatics

Awareness of Bioinformatics			
		Yes	No
Bioinformatics	135	135	
		100%	0%

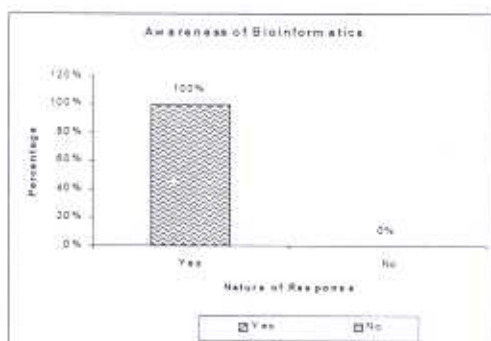


Fig 1: Awareness of Bioinformatics

### Bioinformatics is?

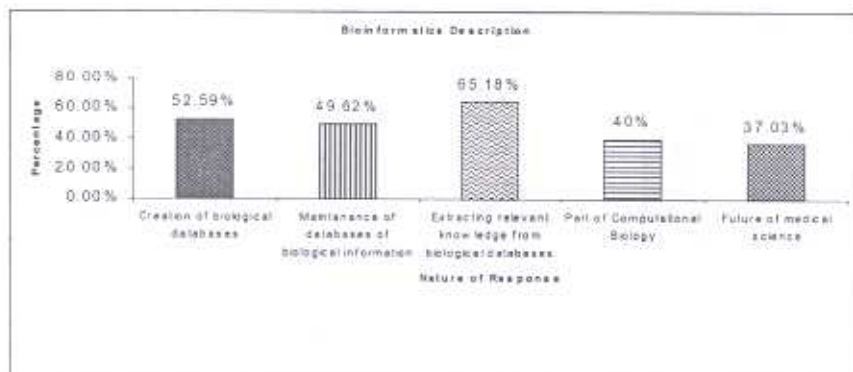


Fig 2: Bioinformatics as defined by students

## Source of information of bioinformatics

Table 3: Source of Information on Bioinformatics

		Fellow students	Classroom information	Reading literature	Library	Internet searching	Commercial advertising
Bioinformatics	135	70 51.85%	9 6.67%	17 12.59%	12 8.89%	17 12.59%	10 7.41%



Fig 4: Percentage source of information about bioinformatics

## Awareness of Biological databases

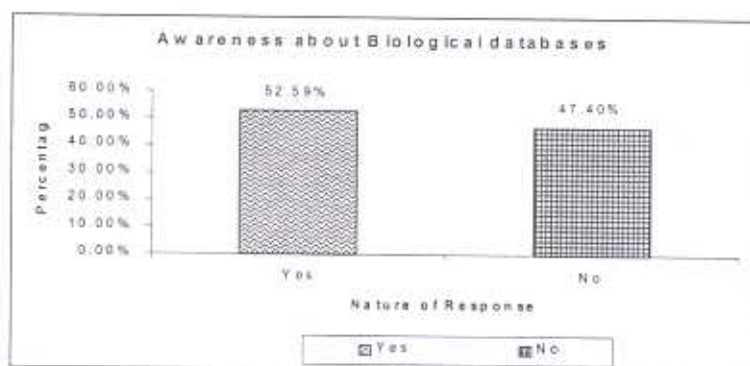


Fig 5: Percentage of awareness about Biological databases

## Source of Information about Biological Database

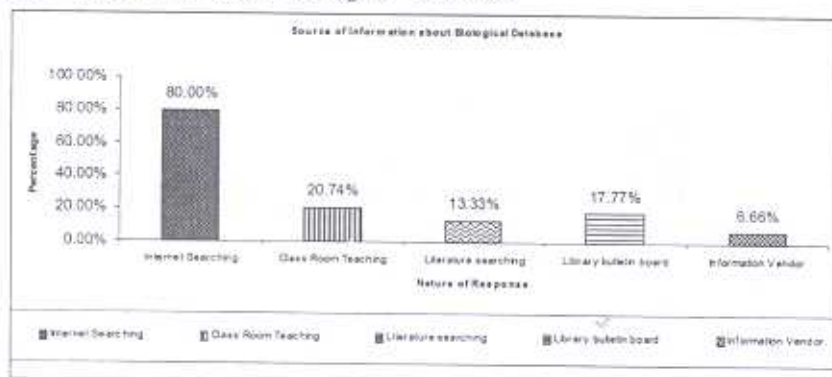


Fig 6: Source of Information about Biological Databases

### Awareness about nucleic acid, protein database

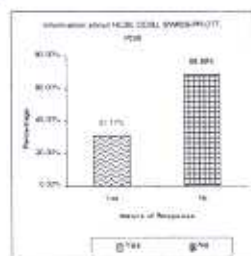


Fig 7: Percentage of awareness about Nucleic Acid, Protein Databases

### Information about NCBI, DDBJ, SWISS-PROT, PDB

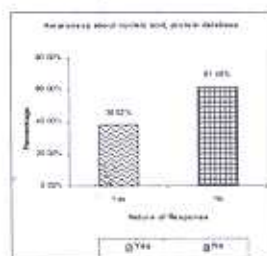


Fig 8: Percentage of awareness about NCBI, DDBJ, SWISS-PROT, PDB.

### Awareness about the URL of NCBI, DDBJ, PDB etc

Table 4: Percentage of awareness about URL of NCBI, DDBJ, PDB etc.

		Yes	No
Bioinformatics	135	39	96
		28.89%	71.11%

### Reaching the specific website

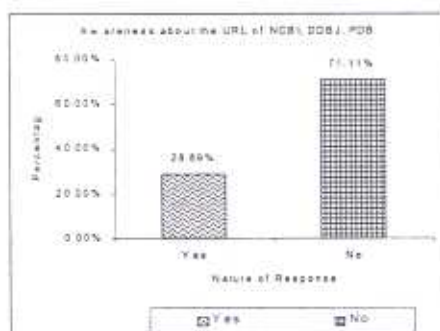


Fig 9: Awareness about the URL of NCBI, DDBJ, PDB etc

### Development of Model

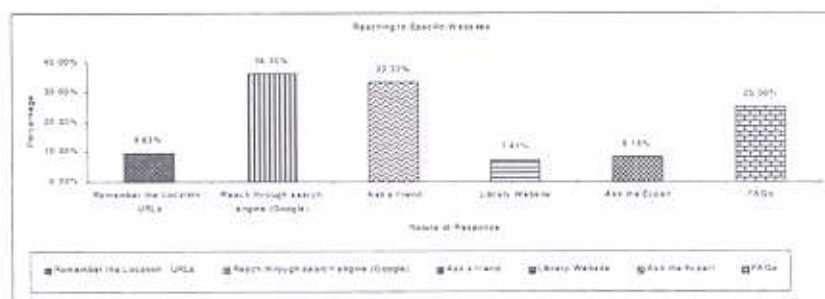


Fig 10: Percentage different approaches of reaching specific websites



By analyzing the data of the questionnaire it can be inferred that the percentage of awareness varies for different set of questions among the student of bioinformatics at JUIT. These variations and difference in the response of the student help to develop the model of bioinformatics digital library. Some of the basic divisions of the bioinformatics which are being used to develop model are enumerated below.

- (i) Bioinformatics databases
- (ii) Bioinformatics institutions
- (iii) Major aggregators
- (iv) Online/offline courses
- (v) Career information
- (vi) Electronic Sources of reading materials
- (vii) Software Tools on Bioinformatics
- (viii) Web based information system in bioinformatics
- (ix) Search guide, a decision support tool, or training materials.
- (x) User education program.

Based on the above taxonomic classification a model was developed using PHP, My SQL, and FrontPage. The finding of the above helped to choose the different set of information to be incorporated to make aware the students. For example majority of the students (36.30%) reach specific website with the help of search engine (Google). Thus we can say that the students are not aware of specific institution program and services pertaining to the bioinformatics. The developed model provides a single access platform to get all the information at single click without wasting the time. Student will be able to search, post and download the desired information and freely navigate from one part to another in order to locate their desired information.

**Bioinformatics Digital Library**  
(Jawahar Institute of Information Technology, JUIT)

Search the Web

Home | Digital Library | Bio Informatics | Program & Research | e-Resources | Online BI Tools | Consulting Facilities

**Administrative**  
Administrator ID:   
Password:

**New User**

**Existing user**

**Biological Databases**

- Browse Bioinformatics
- Digital Library
- Green Catalogue

**Bio Informatics**

- About
- Genomics
- Proteomics
- Drug Design
- Biological Computation
- Molecular Modeling
- Cheminformatics

**Programs & Research**

- Courses
- Scope
- Bioinformatics in India
- Bioinformatics in IIS
- Academician Companies

**e-Resources**

- eBooks
- eJournals
- Patents

**Online BI Tools**

- Similarity & Homology
- Sequence Analysis
- Protein Functional Analysis
- Structure Analysis
- Water Finder
- Phylogenetic Relationship

**Biological Databases**

- Protein Sequence Database
- Nucleotide Sequence Database
- Protein Database
- Literature Database
- Proteomics Database

**Computing Facilities**

- All JUIT
- How to Play in BI
- Web Resources
- Software
- Hardware

**JUIT**  
विद्या तस्य ज्योतिषम्

Fig. 11

## Conclusion

The science of digital library is concerned with the representation, organization, manipulation, distribution, maintenance and use of information, particularly in the digital form. The functional aspect of the digital library is the storage, representation, retrieval and distribution of the data. Bioinformatics digital library at JUIT is an endeavor to provide a centralized access for existing resources in bioinformatics. It can be useful in providing information to develop various analytical tools, such as comparing sequences, 3D modeling of proteins and metabolites, gene sequencing, literature searching, literature citations etc. It also provide platform for the users to access Protein Sequence Database, Nucleotide Sequence Database, Protein Database, Literature Databases, and Proteomic Database for the purpose of curriculum supplement.

## Acknowledgement

I wish to acknowledge my sincere gratitude to Dr Harish Chandra, Librarian, Central Library, Indian Institute of Technology, (Madras), Chennai for his painstaking efforts while guiding me for the Project 'Design and Development of Digital Library in Bioinformatics: a model for universities in Himachal Pradesh: a case study of Jaypee University of Information Technology, Solan (HP)'. His learned counsel and efforts made me to complete my M Phil Thesis of the said project.

## References

1. Baxevanis, A. D. (2000). "The Molecular Biology Database Collection: an online compilation of relevant database resources." *Nucleic Acids Research*, 28(1), 1-7.
2. Collins, F. S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., & Walters, L. (1998). "New goals for the U.S. Human Genome Project: 1998-2003." *Science*, 282(5389), 682-9.
3. Curtis, K. L., Weller, A. C., & Hurd, J. M. (1993). "Information-seeking behaviour: a survey of health sciences faculty use of indexes and databases." *Bulletin of the Medical Library Association*, 81(4), 383-392.
4. Curtis, K. L., Weller, A. C., & Hurd, J. M. (1997). "Information-seeking behavior of health sciences faculty: the impact of new information technologies." *Bulletin of the Medical Library Association*, 85(4), 402-10.
5. Dillon, M. (1981). "Serving the information needs of scientific research." *Special Libraries*, 72, 215-23.
6. French, B. A. (1990). "User needs and library services in agricultural sciences." *Library Trends*, 38, 415-41.
7. Garvey, W. D., Tomita, K., & Woolf, P. (1974). "The dynamic scientific information user." *Information Storage and Retrieval*, 10, 115-31.
8. Grefsheim, S., Franklin, J., & Cunningham, D. (1991). "Biotechnology awareness study, part I: where scientists get their information." *Bulletin of the Medical Library Association*, 79(1), 36-44.
9. Missener, S. and Krawetz, S.A. *Methods in Molecular Biology*, vol 132. Humana Press. 2004.
10. National Human Genome Research Institute. (2000). "Human Genome Project Goals: 1998-2003." Bethesda, MD: Retrieved March 24, 2000 from the World Wide Web: <http://www.nhgri.nih.gov/HGP/>.
11. *Nucleic Acid Research Database Issue*, 2005
12. Palmer, J. (1991a). "Scientists and information: I. Using cluster analysis to identify information style." *Journal of Documentation*, 47(2), 105-29.
13. Palmer, J. (1991b). "Scientists and information: II. Personal factors in information behavior." *Journal of Documentation*, 47(3), 254-75.
14. Persidis, A. (1999). "Bioinformatics." *Nature Biotechnology*, 17, 828-30.
15. PUBMED: [www.ncbi.nlm.gov/entrez](http://www.ncbi.nlm.gov/entrez).

16. Rolinson, J., Meadows, A. J., & Smith, H. (1995). "Use of information technology by biological researchers." *Journal of Information Science*, 21(2), 133-9.
17. Shri Ram. Design and development of bioinformatics digital library: A model for universities in Himachal Pradesh: a case study of Jaypee University of Information Technology. M Phil Thesis Submitted to Alagappa University, 2005.
18. Bartlett, Joan C. (2001) "Facilitating access to and use of bioinformatics resources". *Information Research*, 6(2) Available at: <http://InformationR.net/ir/6-2/wsl.html>
19. Skelton, B. (1973). "Scientists and social scientists as information users: a comparison of results of science user studies with the investigation into information requirements of the social sciences." *Journal of Librarianship*, 5(2), 138-56.
20. Wadman, M. (1999). "Human Genome Project aims to finish 'working draft' next year." *Nature*, 398, 177.
21. Wekipedia: [www.wikipedia.com](http://www.wikipedia.com)
22. Yarfitz, S., & Ketchell, D. S. (2000). "A library-based bioinformatics services program." *Bulletin of the Medical Library Association*, 88(1), 36-48.
23. Harish Chandra, Development and Use of Web based Information Resources with Specific Reference to E-Books: A Case for S&T Libraries. First International Convention on Mapping Technology on Libraries and People Ed by T.A.V. Murthy etal. Ahmedabad, INFLIBNET, 13-15 February, 2003. p. 575-586.