Bioinformatics Digital Library

Shri Ram, Librarian
Jaypee University of Information Technology, Waknaghat, (HP)
Ph: 01792 239 242, 09418109763 e-Mail: shri_ram2576@yahoo.com

Sanjay Kataria, Librarian
Jaypee Institute of Information Technology (Deemed University) Noida
(UP), Ph: 09810503341 e-Mail: katariasanjay2001@yahoo.com

Abstract: Digitization has revolutionized every field of our lives. Starting from the print media it has gone through the broadcast media and now a number of other storage media. The academic learning resources, though have taken lead, yet other such institutions are not lacking behind. Bioinformatics is such an area where the possibilities of digitization can be explored. The paper opens with the remarks that digital libraries are multifaceted systems. The area it has covered includes molecular biology, protein databanks, genome data, computer science and mathematics etc. The new phrase computational biology has been added for the reader's quest. The paper deals in some basic, technical and analytical questions relating to digitization. The application of Internet has been adjusted as the backbone of digitization. The paper also discusses the different field of bioinformatics which has been covered in the development of bioinformatics digital library at Jaypee University of Information Technology, Solan (HP).

Key Words: Digital Library, Bioinformatics Digital Library, Biological Databases.

INTRODUCTION

The knowledge in the present day scenario is changing rapidly in multidimensional ways. New developments in the science and technology environment leading to enormous growth of literature and information resources. Due to change in the information needs of the scientific society, the scientific society is dependent not only upon print information resource but also on electronic information resources, databases, CD-ROM etc. To keep the user

aware of current developments in their respective areas, it is necessary for information science to synthesize and integrate new ideas at single platform for common access. This will enable the users to enhance their academic and research skills in their fields at minimum expanse of time, money and matters. As proposed by one of the five laws of library science given by Dr. S. R. Rangnathan, 'save the time of reader'. This change brings the role of digital library.

Digital Library Federation (DLF) defines digital library as 'Digital Libraries are the organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for the use by a defined community or set of communities'. The digital libraries are a large varied set of services that include the service to support the management of collection, to provide reliable storage, to aid queries, formulation and execution etc. In today's scenario the information flow requires organized collection, management and presentation of the information for the easy access by end user at minimum time.

In recent past the technological shift has been observed from the traditional libraries to the digital environment due to the advancement in Information and Communication Technology (ICT). Internet based services have revolutionized the science and technology activities in the world. In present day changing society the digital library technology can link multiples resources; establish link between various digital libraries and information centers and finally sources to end users in transparent ways. Universal access of the information through digital technologies is not only limited to document surrogate, but they extend to digital artifacts that cannot be represented or distributed in printed form. Digital libraries are viewed as a system which provides coherent access of to large and organized repositories of information and knowledge to a community of users.

BIOINFORMATICS

Bioinformatics is a science which developed from integration of biology and Information Technology (IT). It is a multidisciplinary field comprising of molecular biology, genetics, and Mathematics and computer science. Today, Bioinformaticians are in great demand in academia as well as in industry. The major goal of the bioinformatics is storage and meaningful analysis of raw biological data knowledge discovery. Researchers, students and faculty working in bioinformatics regularly seek information from the library resources in bioinformatics research. For example, they collect literature, experimental information from the following sources

(a) NCBI, Enterez PubMed online databases, OMIM etc-**Molecular Biology**

(b) **Protein Data banks (PDB)** – for proteins related information.

(c) **PubMed and Medline**: Enables keyword search for relevant literature retrieved from various journals

AREAS IN BIOINFORMATICS

Bioinformatics has evolved into full-flagged scientific discipline over the last decade. The definition of bioinformatics is not restricted to computational molecular biology or computational structural biology. It now encompasses fields such as comparative genomics, structural genomics, transcripteomics, proteomics, metabolic pathways engineering as well as computing and mathematical techniques. Developments in these fields have direct implications to health care, medicine, discovery of new next generation drugs, agricultural products, environmental protection etc. Information resources from these linked fields can be searched in library and information centers in three ways.

1. Basic questions: locating resources, programs and databases
2. Technical questions: Identifying appropriate tools for academic and research needs
3. Analytical Questions: planning experiments or in depth assistance with data analysis.

It means the library and information professionals can answer the above mentioned basic questions. Digital library aims to extend support to expand the services to bioinformatics development at the single platform. It can also provides services like the searching of primary literature; genomics sequence databases, data and knowledge management and communication. These integrated services provided by bioinformatics digital library can enhance the skills of bioinformatics professional significantly.

INFORMATION SOURCE ON BIOINFORMATICS

Plethora of information are available in the field of bioinformatics, most of these information available are related to new techniques such as drug development, software tools for biological data analysis, drug designing, technology introduced in bioinformatics and database information. Established ideas and methods are constantly being refined and are appearing in journals, books, software packages, and through internet technology or web documents and published documents. In With the intervention of Internet and computer network models, the distribution of information became faster and efficient. The bioinformatics information is now available throughout the world to the scientific organizations with the help of Internet. National Medical Libraries are the prime sources of information for biomedical researchers. The science of informatics is concerned with the representation, organization, manipulation, distribution, maintenance and use of information, particularly in the digital form. Considering the requirement of platitude for accessing the vast bioinformatics resources, a model system of digital library in bioinformatics at Jaypee University of Information technology has been developed. This library is currently functional through Intranet. (172.6.73.16)

While developing the bioinformatics digital library at Jaypee University of Information Technology (JUIT), Solan (HP) we have analyzed the following category of Digital Information System.

(i).   A digital library collection that includes the "digital original" resources, which are sometimes referred to as resources which are "born digitally".

(ii) Digital Library that comprises "digital surrogates", which are created from traditional information resources through format conversion.

Above mentioned both type of digital library system has analyzed, data collected and put at a single system "Bioinformatics Digital Library" under following headings with the help of PHP Computer Programming Language, My SQL query language, DHTML, FrontPage and HTML.

(a) **Digital Library**: Digital Surrogate Created from the traditional information sources.

(b) **Bioinformatics**: The Information related to bioinformatics, genomics, proteomics, drug design, biological computation, molecular modeling, and Chemo-informatics.

(c) **Program and Research**: Information about institute running courses and programs in various disciplines and career opportunity in bioinformatics.

(d) **E-Resources**: comprises link to internet resources of electronic journals, electronic books, and e-patents information. Some of the important digitally accessible e-journals are included in this section are:

   a) **Genome Research**: http://www.genome.org/

   b) **Protein Science:** http://www.proteinscience.org/

   c) **RNA:** http://rnajournal.org/

   d) **Bioinformatics:** http://bioinformatics.oxfordjournals.org/

   e) **Nucleic Acid Research:** http://nar.oxfordjournals.org/ and so on. About more than 100 such open access journals links has been provided to access the scholarly peer reviewed journals.

(e) **Biological Database Information:** The bioinformatics more dependent on the biological database information to carry out R&D activities. These databases are related to nucleic acid information, proteins, genomics and proteomic information. The various organizations and institutes which are providing store house to these information for the global access to databases are given link in this section. Some of the such important centers are:

    a) **Protein Information Resources:** Located at Georgetown University Medical Center. http://nbrf.georgetown.edu

    b) **Nucleic Acid Database:** Sequence Retrieval Systems (SRS). http://www.srs.ebi.ac.uk

    c) **EMBL:** http://www.ebi.ac.uk

    d) **DNA Databank of Japan:** http://ddbj.nig.ac.jp/

    e) **Protein Sequence Data Bank:** Swiss-Prot. http://wwwebi.ac.uk/swissprot etc.

(f) **Literature Databases:** The literature resources play and important role in R&D activities in any of field. The some of the important literature sources on bioinformatics are:

    a) **ENTERZ**: Enterez is search engine for life science literature. It is text based search and retrieval systems used at NCBI for various services including PubMed, nucleotide and protein sequences, protein structures complete genomes, taxonomy OMIM and many others.

    b) **PubMed**: It is available via NCBI Enterez Retrieval Systems developed at NLM, USA

c) **MEDLINE**: Medical Literature Analysis and Retrieval Systems Online is the primary component of PubMed providing literature services of NCBI.
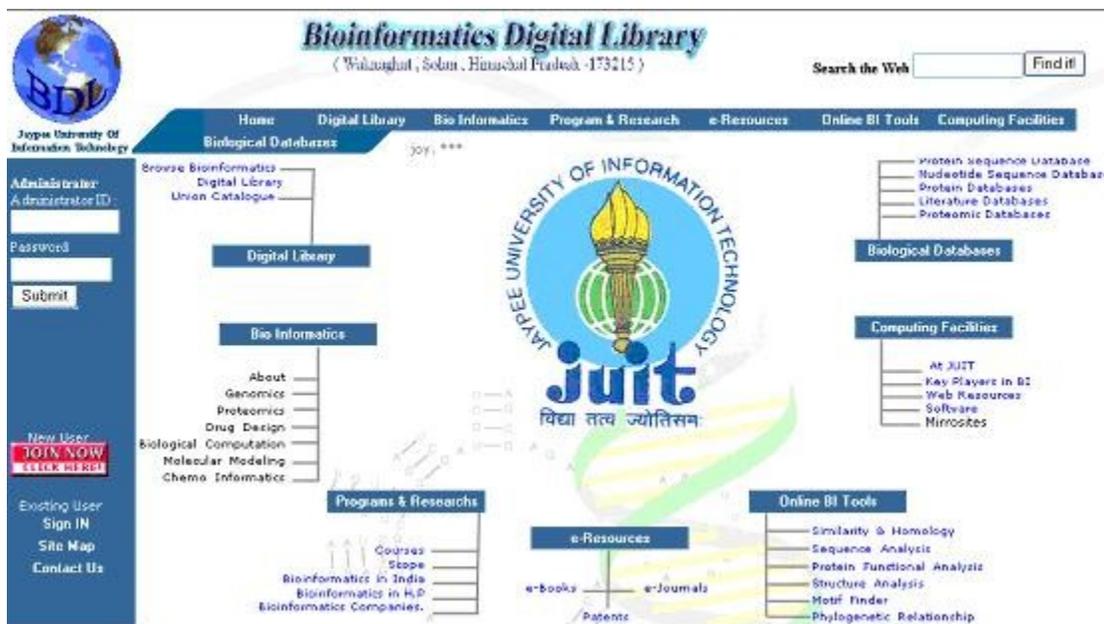

Fig 1: Bioinformatics Digital Library at JUIT

Conclusion

The science of informatics is concerned with the representation, organization, manipulation, distribution, maintenance and use of information, particularly in the digital form. The functional aspect of the informatics 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 etc.

It also provide platform for the users to access Protein Sequence Database, Nucleotide Sequence Database, Protein Database, Literature Databases, and Proteomic Database.

REFERENCES

1. **Alpi, Kristine**. Bioinformatics training by librarians and for librarians; developing the skills needed to support molecular and clinical genetics information instructions.

2. **Baldi Pierre** and **Brunak, Soren**. Bioinformatics: Machine learning approach. 2nd ed. MIT Press, Cambridge. 2003.

3. **Bartlett, Jean C**. Facilitating access to and use of bioinformatics resources. Inforamation Research. 6(2). 2001.

4. **Chandra, Harish**. Development of web based library and information support services: a foundation stone for the building digital libraries. Lecture delivered at National Workshop on building Digital Libraries on June 17-18, 2005 Jaypee university of Information Technology, Noida.

5. **Dubey, Christopher**, et al.. Developing bioinformatics training: bridging gap between computer science and biomedicine. Proceedings of the AMIA 2002 annual symposium. P220-224. 2002.

6. **Grefshem, Suzane**. Biotechnology awareness study, Part- 1: where scientist get their information. Bulletin of Medical Library Association. 79(1). 1991.

7. **Guise, Nonzia Bellinsoli**. Integrating health science librarians into biomedicine. Bulletin of Medical Library Association. 84(4). P534-540. 1996

8. **Hersh, William**. Bioinformatics digital library. JCDL'02. July 13-17, 2002.

9. **McCray, Alexa T**. and **Gallager, Marie E**. Principles for digital library development. Communication of the ACM. 44(5). P49-54. 2001.

10. **Ranganathan, S R**. Laws of Library Sceince. Allied Publishers. New Delhi.

11. **Shri Ram**. Design and development of digital library in bioinformatics: a model for universities in Himachal Pradesh: a case study of Jaypee University of Information Technology. M. Phil. (Library and Information Science) Theses. Alagappa University. 2005.

12. **Singh, Sukhdev**, **Gaba, Surinder K**. and **Pandita, Naina**. Architecture and building medical digital library at NIC of India. What exhists and what required for the MedLib@NIC?. Proceedings of the International Conference on Digital Library, New Delhi (India).1995.

13. **Vladimir, B. Bajic** et al.. From Informatics to Bioinformatics. Proceedings of the First Asia-Pacific Bioinformatics conference on Bioinformatics, January 2003. p3-12. 2003.

14. **Waters, D.J**.. What are digital libraries? *CLIR Issues*, July/August. URL: http://www.clir.org/pubs/issues/issues04.HTML. 1998.

15. **Yarfitz, S.** and **Ketchell, D. S**. A Library based bioinformatics service program. Bulletin of Medical Library Association. 88(1). P36-48. 2000.