BIOINFORMATICS

Bojana Boh
Faculty of Natural Sciences and Engineering, International Centre for Chemical Studies, University of Ljubljana, Slovenia

Keywords: Bioinformatics, genomics, proteomics, biosciences, Internet, online, bibliographic databases, value added processing, information density, factual databases, nucleic acids, proteins, sequencing, and structure elucidation

Contents

Summary

Bioinformatics, a specialized sub-branch of scientific and technical informatics, interlinks computer sciences with genetics, biotechnology, biochemistry, medicine, biology and pharmaceutical sciences. The number of databases in biosciences and their file sizes has been in rapid expansion. The largest bibliographic databases, accessible online, via Internet or on CD-ROMs, contain several million records. Due to the achievements in protein and nucleic acid sequencing and structure elucidation, new generation factual databases underwent an exponential growth in the second half of 1990s and in the 2000s. Internet and the World Wide Web with the introduction of hypermedia-based information systems had a profound impact on the development and use of databases. Bioinformatics has being actively supporting all biosciences, including nucleic acid, peptide and protein sequencing, genome and chromosome mapping, and three-dimensional structure elucidation and expression research. On higher levels of information processing, algorithms and probability theories have been applied for sequence data analyses, determination of gene functions, recognition of phylogenetic and evolutionary relationships, predictions of structure-activity relationships, and for the target-oriented synthesis of biomolecules.

1. Introduction

Accelerated scientific research, rapid technological developments, and competitiveness in national and global markets have become the main driving forces for the intensive

growth of information during the last decades. The success of research, development and marketing thus depends to a large extent on fast and reliable transfer of information and on the ability to organize information into knowledge. Combined with the achievements in telecommunications and the computer industry, an ever increasing amount of information resulted in the development of informatics as a scientific discipline, and in the prosperity of the information industry, which collects, organizes, processes and sells data and information. Informatics, at the beginning a uniform scientific discipline, has become specialized and linked to different scientific, technological and commercial fields. Examples of such diversification include business informatics, scientific and technological informatics, and bioinformatics with genomics and proteomics.

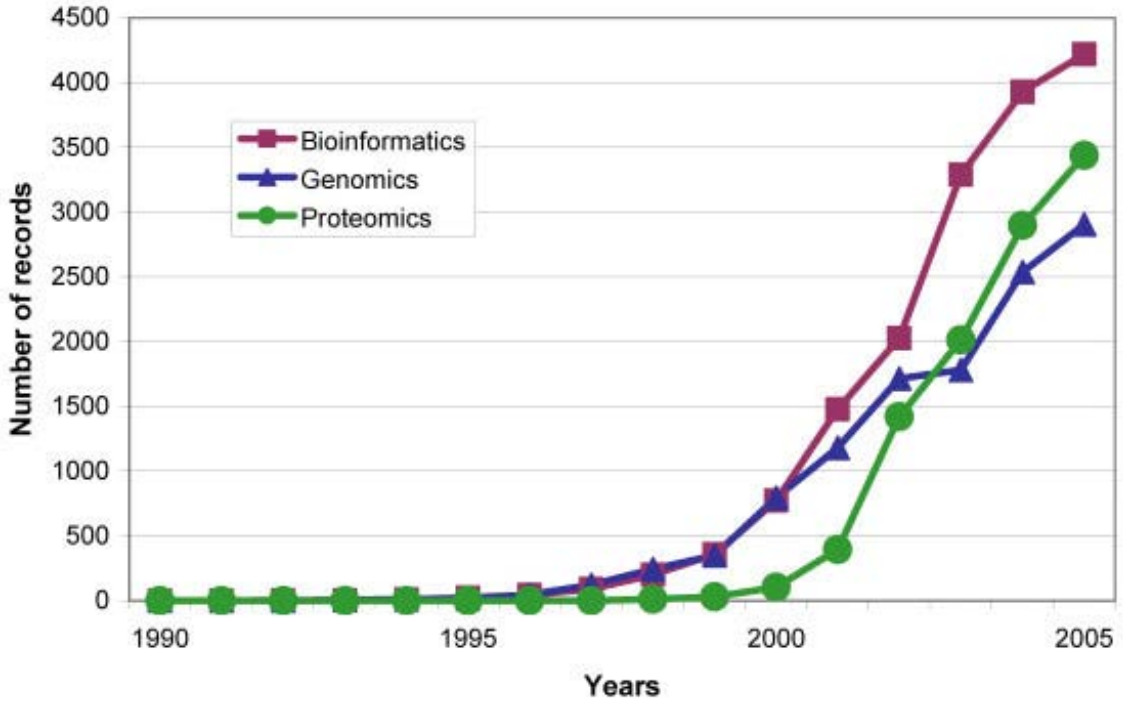

Figure 1: Number of new documents on bioinformatics, genomics and proteomics in the CAPlus database.

Bioinformatics, bridging the life and computer sciences, has to follow radical changes in modern genetics **(**see also - *Genetics and Molecular Biology*), biotechnology, biochemistry **(**see also - *Biochemistry***)** and pharmaceutical science, where an intensive growth of information is particularly evident, both in the theoretical and the commercial domains. In addition to its traditional functions related to building and searching of bibliographic, factual and relational databases, molecular bioinformatics as a newly emerging interdisciplinary research area comprises the development and application of algorithms for the analysis, interpretation and prediction of data in the biosciences. The application of algorithms from computer sciences, including artificial intelligence, machine learning, genetic programming, evolutionary algorithms and neural networks, to molecular biology, especially DNA and RNA sequence analysis and protein engineering **(**see also - *Protein engineering***)**, has become a routine in bio-scientific research. Bioinformatics and its new branches - genomics and proteomics - thus brought

software systems and technologies with computational methods to the analysis and processing of molecular structure, and genomic and proteomic data, and linked them with new types of data collections, such as sequence databases, metabolic pathways, phenotypes, variety collections, gene expression atlases, and phylogenetic trees.

2. Levels of information processing

Computerized scientific and technological information can provide support for both, memory and logic, depending on the level of information processing.

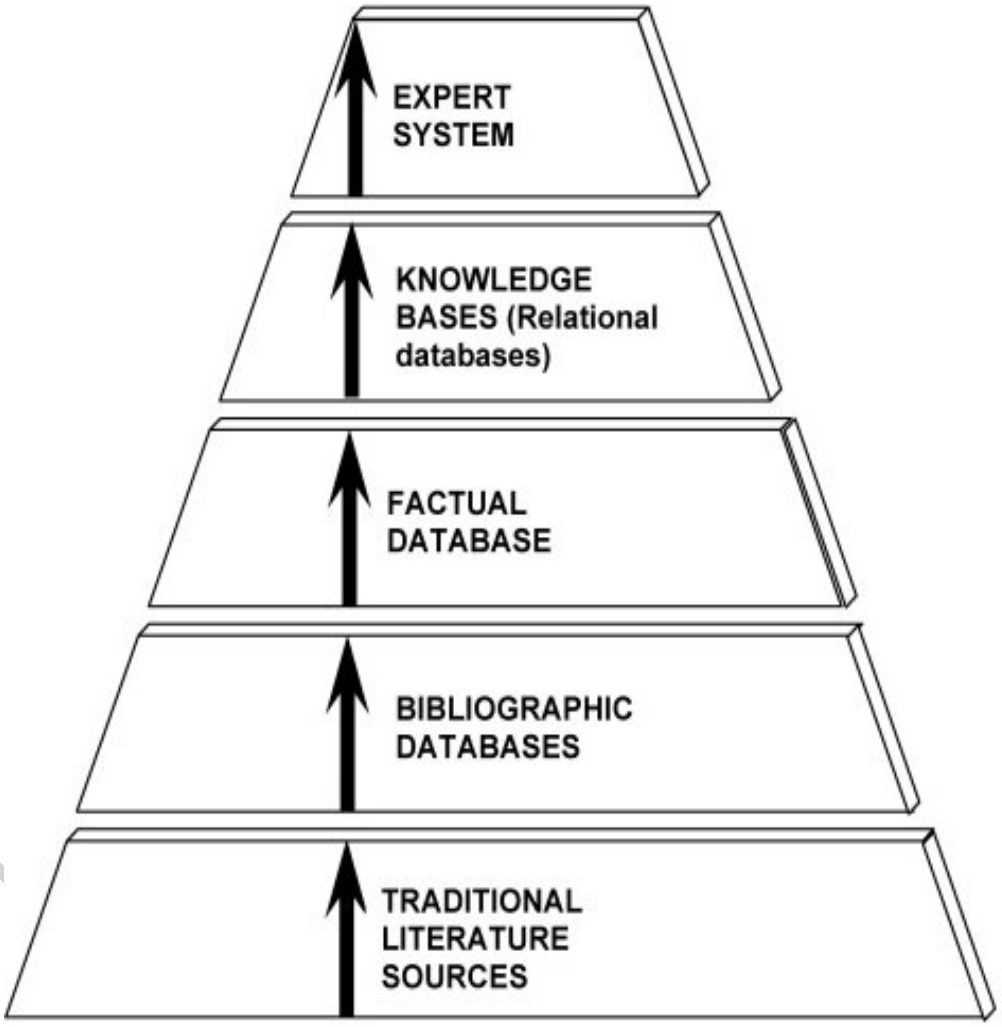

Figure 2: Levels of information processing

-
-
-

Bibliography

Attwood T.K. and Parry-Smith D.J. (1999). *Introduction to Bioinformatics*, 218 pp. Essex: Addison Wesley Longman Ltd. [This book systematically introduces the essential aspects of bioinformatics]

Boh B. (1996). Organisation of Biotechnological Information into Knowledge. *World Journal of Microbiology and Biotechnology*, **12**(5), 425-437. [This paper gives further examples of information methods and information-based research strategies]

Kardos D. and Boh B. (2000), An Information Method for Achieving Value-Added Processing of Bibliographic Databases in Science and Technology", *Online Information Review*, **24**(4), 294-301. [This paper describes the methodology of value-added database processing for the recognition trends in research and development]

Kornhauser A. (1989). Searching for Patterns of Knowledge in Science Education. *New Information Technologies in Higher Education: Studies on the Introduction of New Information Technologies in Higher Education in the Europe Region*, (eds. Calude, C., Chitoran, D. and Malitza M.), 155-168. Bucharest: European Centre for Higher Education - CEPES. [This paper describes information methods and approaches in scientific data processing]

Kornhauser A. and Boh B. (1992). Information Support for Research and Development in Biotechnological Applications. *Biotechnology: Economic and Social Aspects*, (eds. DaSilva E.J., Ratledge C. and Sasson, A.), 309-353. Cambridge: Cambridge University Press. [This chapter presents and discusses information sources, especially databases, for biotechnological research and development]

STN International. *Your Connection to Science and Technology*, http://www.stn-international.de/ (accessed July 2006). [This is the Internet homepage of the main host of scientific and technological databases. Summary sheets with full descriptions of databases are available through this site]

STN on the Web. *Free Search Preview*, http://stnweb.cas.org/ (accessed July 2006). [This is the Internet access to STN database clusters and individual databases. For a given search profile, the number of hits in a database can be viewed free of charge]

Biographical Sketch

Bojana Boh was born on 8 March 1960 at Postojna, Slovenia. *Education:* B. Sc. Biology/Biochemistry, 1983, Biotechnical Faculty, University of Ljubljana, Slovenia; M.Sc. Biochemistry, 1986, Biotechnical Faculty, University of Ljubljana, Slovenia; Ph.D., Chemistry, (Combination of Information and Laboratory Methods in Microencapsulation of Proteins by Interfacial Polymerization), 1991, Faculty of Science and Technology, University of Ljubljana, Slovenia. *Post-Doctoral Research*: 1993—Institute of Pharmaceutical Technology, Johann-Wolfgang-Goethe University, Frankfurt am Main, Germany. *Employment:* University of Ljubljana, Faculty of Natural Sciences and Engineering, Department of Chemical Education and Informatics/International Center for Chemical Studies, Researcher, 1983–1992., Assistant Professor 1992-2000 and Associate Professor since 2000. *Fields of Work:* (1) Graduate and post-graduate educational programs including chemistry of natural products, scientific and technological informatics, bioinformatics. (2) Basic research and R&D projects: development of information systems and information methodology with applications in chemistry, biochemistry and biotechnology; research in the field of natural products and their derivatives (*Ganoderma* triterpenoids and polysaccharides, natural dyes, essential oils); development of microencapsulation technologies for applications in agriculture, food, pharmaceutical and paper industries; (3) International projects: (a) UNESCO programs: MIRCEN-BITES (Microbial Resources Centres - Biotehnological Information Exchange System); university-industry cooperation; structuring of scientific and technical data for the development of specialized information systems, (b) EU Tempus-Phare project on University-Industry Cooperation - Teaching Strategies; (c) EU COST projects on Bio-microencapsulation; (d) The World Bank projects: Environmental Education programs (1990-93); Scientific Information and Literature (1993); Interactive Educational Technologies (1994); Higher Education and Research (1999-2000), (e) a bilateral project with the University of Nairobi, Kenya, on cultivation of *Matricaria chamomilla* and extraction of essential oils. *Publications:* Active participation as an author or co-author of 8 monographs, 87 scientific articles, 2 patents, 160 reports for the industry, 11 databases; active participation in about 80 scientific

meetings. *Accomplishments:* Pro Natura (Slovenia) / Boehringer Ingelheim Fond (Germany) award for young researchers, 1992.