

Library Collection Development for Professional Programs: Trends and Best Practices

Sara Holder
McGill University, Canada

A volume in the Advances in Library and
Information Science (ALIS) Book Series



Managing Director:	Lindsay Johnston
Senior Editorial Director:	Heather A. Probst
Book Production Manager:	Sean Woznicki
Development Manager:	Joel Gamon
Assistant Acquisitions Editor:	Kayla Wolfe
Typesetter:	Jennifer Romanchak
Cover Design:	Nick Newcomer

Published in the United States of America by
 Information Science Reference (an imprint of IGI Global)
 701 E. Chocolate Avenue
 Hershey PA 17033
 Tel: 717-533-8845
 Fax: 717-533-8661
 E-mail: cust@igi-global.com
 Web site: <http://www.igi-global.com>

Copyright © 2013 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Library collection development for professional programs: trends and best practices / Sara Holder, editor.
 pages cm

Summary: "This book addresses the challenging task of collection development in modern academic libraries and contains practical advice and innovative strategies for current collection development librarians and future librarians seeking guidance in this complex position"--Provided by publisher. Includes bibliographical references and index.

ISBN 978-1-4666-1897-8 (hardcover) -- ISBN 978-1-4666-1898-5 (ebook) (print) -- ISBN 978-1-4666-1899-2 (print & perpetual access) (print) 1. Academic libraries--Collection development--United States--Case studies. 2. Academic libraries--Collection development--Canada--Case studies. 3. Collection development (Libraries) I. Holder, Sara, 1967-
 Z675.U5L5185 2013
 025.2'1877--dc23

2012009488

This book is published in the IGI Global book series Advances in Library and Information Science (ALIS) Book Series (ISSN: 2326-4136; eISSN: 2326-4144)

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Chapter 16

Developing a Library Collection in Bioinformatics: Support for an Evolving Profession

Victoria Martin
George Mason University, USA

ABSTRACT

This chapter provides guidelines for developing a university library collection for bioinformatics programs. The chapter discusses current research and scholarly communication trends in bioinformatics and their impact on information needs and information seeking behavior of bioinformaticians and, consequently, on collection development. It also discusses the criteria for making collection development decisions that are largely influenced by the interdisciplinary nature of the field. The types of information resources most frequently used by bioinformaticians are described, specific resources are suggested, and creative options aimed at finding ways for a bioinformatics library collection to expand in the digital era are explored. The author draws on literature in bioinformatics and the library and information sciences as well as on her ten years of experience providing bioinformatics user services at George Mason University. The chapter is geared towards practicing librarians who are charged with developing a collection for bioinformatics academic programs as well as future librarians taking courses on collection development and academic librarianship.

INTRODUCTION

To date, little research has been conducted on developing library collections in bioinformatics. The author aspires to fill this gap by offering a broad summary of collection development principles and practices as seen from the perspective of a liaison librarian for whom collection development is just

one component of her job. The uniqueness of such a perspective lies in the fact that liaison librarians bring to collection development decisions not only their subject expertise but also their knowledge of the specialized needs of faculty and students within their assigned departments. Liaison librarians are also in a good position to involve faculty with collection development decisions to ensure

DOI: 10.4018/978-1-4666-1897-8.ch016

that the needs of all faculty and students are being met. Because of their direct experience with library patrons, liaison librarians are also able to provide insight into the information seeking patterns of the users that will assist in building a more balanced and meaningful collection.

The first section of this chapter describes current trends in bioinformatics study, teaching, and research that span several disciplines and are largely driven by new technologies and innovative research methods. It also examines how these trends affect the information seeking patterns of bioinformaticians who rely on timely access to current research findings. Understanding of these patterns allows a librarian to formulate the best approaches for building a collection in bioinformatics.

The next section describes the characteristics of a bioinformatics library collection, identifies the types of resources used most frequently by bioinformatics faculty and students, and suggests selection criteria of various library materials. Alternative scholarly communication models, such as open access, are discussed. Skills and competencies and educational and training opportunities that will help a bioinformatics librarian achieve greater success are described. The final section emphasizes that collection development in bioinformatics is a challenging but exciting task that is mainly learned on the job where true mastery comes with the experience of learning and doing. Specific resources are suggested in a Key Resources section at the end of this chapter.

BACKGROUND

Definition of Bioinformatics

There is no unified or comprehensive definition of bioinformatics in the existing journal and reference literature (Altman & Mooney, 2006; Buehler & Rashidi, 2005; Fenstermacher, 2005; Hancock & Zvelebil, 2004; Rédei, 2008; Tramontano,

2009) and the perception of what bioinformatics means is still evolving. The challenge of defining bioinformatics is due to the constantly changing scope of the discipline, which encompasses many areas of study such as DNA sequence alignment and the analysis of protein structure and function, as well as to the blurring boundaries between bioinformatics and emerging research fields such as neuroinformatics and ecoinformatics. Although the definition of bioinformatics is still the matter of some debate, the underlying idea in these various definitions of this term is the overlapping of two main fields—biology (the study of life) and informatics (the science of processing data for storage and retrieval), hence giving rise to the name bioinformatics. For the purpose of this chapter, bioinformatics is defined in general terms as the application of computers and computational techniques for the interpretation of biological data.

Overview of the Field

Bioinformatics, sometimes referred to as computational biology or biological information science, is a rapidly growing research and academic field that draws on a range of disciplines, including biology, computer science, information technology, and mathematics (whether bioinformatics and computational biology are the same or distinct is still a matter of debate). Bioinformatics also utilizes computational techniques borrowed from the fields of artificial intelligence, statistics, and biochemistry. The combined strength of scientific collaboration among several disciplines has brought bioinformaticians closer to an understanding of living systems across multiple levels of biological organization and has improved their ability to understand the root causes of human, animal and plant diseases and to find new cures for them.

Though it dates back to the 1960s following the discovery of the DNA double helix, bioinformatics has emerged and gained recognition as a separate research field during the 1990s when

the emergence of high-throughput computational techniques, also known as “omics,” revolutionized the research paradigm of the biosciences. These techniques (genomics, proteomics, and metabolomics, to name a few) have produced such vast quantities of complex biological data that specialized software tools were designed by experts in the field to facilitate the interpretation and analysis of this data. Furthermore, the development of automated DNA sequencing and other innovative research methods has reduced the cost and time needed in discovering the genetic makeup of various organisms. World-wide online molecular biology databases (such as GenBank and Protein Data Bank) were created, offering scientists the opportunity to access a wide variety of biological data.

This confluence of technological and scientific advancements, coupled with the power of the Internet, has enabled researchers to communicate with colleagues more quickly and efficiently and to share biological data with other scientists in a collaborative way, while dramatically improving the productivity of bioinformatics research. It has taken research from a place where it was common to study a biological problem one gene at a time to today's standard where it is possible to model multiple genes working together and performing complex biological functions. It has enabled scientists, in just over a decade, to map the entire human genome.

The Human Genome Project, started in 1990 and completed in 2002, mapped and sequenced the three-billion nucleotide bases of DNA in human cells. The studies of DNA variation continued in the International HapMap Project (completed in 2005), which described the common patterns of human genetic variation. Since the completion of these two projects, a Genome-Wide Association Study (GWAS) was undertaken in 2005 with the purpose of uncovering the associations between genetic variants and human diseases and traits.

These discoveries are making tremendous contributions to humankind such as new drug

design and development, diagnoses of hereditary diseases, and gene therapy and treatments. The impact of bioinformatics research is not just limited to the healthcare field. It has many practical applications in environmental science, biotechnology, and agriculture. Some specific examples of practical applications of this include genetic engineering of drought-and-disease-resistant plants, microbial genome alternations for energy production, environmental cleanup, industrial processing, and waste reduction (Fulekar, 2009).

Financial support from the government has been crucial to the growth of bioinformatics research and training, which is supported through grants from the National Institutes of Health, National Science Foundation, Department of Energy, and other foundations such as Burroughs-Wellcome Fund, Keck Foundation, Howard Hughes Medical Institute, PhRMA Foundation, Sloan Foundation, and Culpeper Foundation (Zatz, 2002).

Key Organizations

The establishment of specialized research institutes and organizations throughout the world has played a key role in the progress of bioinformatics research. In the United States, the National Center for Biotechnology (NCBI) (<http://www.ncbi.nlm.nih.gov/>), a unit of the National Library of Medicine within the National Institutes of Health, provides and maintains nucleotide sequence information on genes and clones and has been a developer of analytical and education tools and databases in the field. NCBI currently hosts 39 literature and molecular biology databases, among which are the databases of canonical importance to biologists: GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), containing nucleotide sequences, the related database RefSeq (<http://www.ncbi.nlm.nih.gov/RefSeq/>), containing sequences and their protein products, OMIM (Online Mendelian Inheritance in Man) (<http://www.ncbi.nlm.nih.gov/omim/>), a catalog of human genetic disorders and their causative genes, and PubMed ([271](http://www.</p></div><div data-bbox=)

ncbi.nlm.nih.gov/pubmed/), which comprises more than 21 million citations for biomedical literature from the database MEDLINE produced by the National Library of Medicine. NCBI is one of the world's major resources for organizing and delivering molecular biology information and its resources are accessed by millions of users from all over the world on a daily basis (Rapp & Wheeler, 2005; Sayers, et al., 2011).

In Europe, the European Bioinformatics Institute (EBI), located in Great Britain, has been an important developer for essential computational biology resources and tools and provider of advanced bioinformatics training to scientists at all levels. EBI hosts one of the world's most important collections of biological data, including DNA sequences (EMBL-Bank) (<http://www.ebi.ac.uk/embl/>), protein sequences (UniProt) (<http://www.ebi.ac.uk/uniprot/>), animal genomes (Ensembl) (<http://useast.ensembl.org/index.html>), three-dimensional structures (the Protein Data Bank in Europe) (<http://www.ebi.ac.uk/pdbe/>), data from gene expression experiments (Array-Express) (<http://www.ebi.ac.uk/arrayexpress/>), and protein-protein interactions (IntAct) (<http://www.ebi.ac.uk/intact/>). EBI also provides free-of-charge, publicly available bioinformatics services such as data submissions processing, biological database production, and analysis and retrieval systems and tools.

Bioinformatics Study and Teaching

Over the last two decades, education in bioinformatics has evolved from informal workshops and training courses to structured degree programs (Cummings & Temple, 2010; Hemminger, Losi, & Bauers, 2005; Ranganathan, 2005; Schneider, et al., 2010). The number of graduate programs, as well as enrollment in these programs, expanded rapidly during this period. At last count, there were 33 universities in the United States that offered graduate degrees in bioinformatics (Peterson's, 2011). The growth in undergraduate bioinformat-

ics programs has been slower than for graduate programs. In general, bioinformatics education for undergraduates has been developed by either introducing new courses with a bioinformatics focus or by incorporating bioinformatics course modules into existing courses and laboratories in the natural sciences (Furge, Stevens-Truss, Moore, & Langeland, 2009; Weisman, 2010). Instruction on the use of major molecular biology databases and tools, such as GenBank, OMIM, and BLAST, has also been introduced within the undergraduate biology curriculum (Dinkelman, 2007).

In contrast to more traditional academic programs, the interdisciplinary nature of bioinformatics brings together students with two different backgrounds—those with a primary background in the biological sciences and those with a primary background in computer science. This presents a challenge for academic institutions whose mission with regard to bioinformatics programs is to train the next generation of researchers with a strong grounding in both biology and computer science. Many universities are responding to this challenge by designing a flexible, integrated curriculum of several disciplines that is tailored to the different backgrounds of incoming students, and by involving faculty from two or more departments (Cummings & Temple, 2010; Schneider, et al., 2010; Zatz, 2002).

The bulk of graduate bioinformatics courses are divided among the following three main categories: fundamental bioinformatics (courses in genomics, proteomics, bioinformatics, and computational biology), biology (courses in biochemistry and molecular and cell biology), and computer science (courses in computer programming and database management). However, some bioinformatics education does not occur within the traditional classroom setting. The past decade has seen a rapid integration of eLearning components in bioinformatics courses. eLearning, which can be defined as the use of electronic media and internet technologies to enhance educational activities, presents instructors with a new opportunity

to complement traditional face-to-face lectures with eLearning tools such as Web-based virtual laboratories, wikis, and tutorial videos (Kawano, Ono, Takagi, & Bono, 2011; Saravanana & Shanmughavel, 2007; Weisman, 2010; Zhang, Zhao, & Nunamaker, 2004).

INFORMATION NEEDS OF BIOINFORMATICIANS

To maximize the use and value of university library collections in bioinformatics, it is essential to know the information needs and information-seeking behavior of bioinformatics faculty and students. Several studies, mainly using surveys and interviews, have characterized bioinformatics user groups and their information needs (Dennie, 2010; Geer, 2006; MacMullen & Denn, 2005; Tennant, 2005; Vaidhyanathan, 2008;). Their findings revealed that two sources of information are of primary importance to bioinformaticians: 1) molecular biology databases, and 2) peer-reviewed journal literature. Other sources of information included books, used primarily for teaching and learning, and so-called “invisible college” resources, the informal methods of scholarly communication, such as blogs, email, discussion forums, and professional and scholarly hubs.

Molecular Biology Databases

Molecular biology databases, most of which are free and open to the public, contain original research data such as genomic sequences, microarray data, biochemical properties, and protein structures. With the rapid accumulation of biological knowledge, these databases provide mechanisms to better store and represent data in a controlled and structured way. They also allow scholars around the world to contribute their own data to these databases as well as to analyze and reuse the data collected by others.

Depending on the type of data contained in molecular biology databases, they can be divided into two types: primary and secondary. Primary databases include the principal output from experimental work. An example of a primary database is ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>) from EBI, which contains microarray expression data. Secondary databases draw upon primary databases and bring together information centered on a particular topic. An example of a secondary database is Mouse Genome Informatics (MGI) (<http://www.informatics.jax.org/>) from the Jackson Laboratory, which curates genomic, genetic, and functional data associated with the laboratory mouse.

NCBI in the United States and EBI in Great Britain are the two main biological database curators (Brooksbank, Cameron, & Thornton, 2010; Chiang, 2004; Sayers, et al., 2011). A collection of molecular biology databases and data mining tools maintained by NCBI and made freely available through the NCBI's website (www.ncbi.nlm.nih.gov) is of particular importance to the field. Bioinformaticians use these databases on a regular basis to obtain primary data such as genomic sequences, biochemical properties and structures, and evolutionary histories of organisms. In addition, many molecular biology databases contain bibliographies as part of their records. Such bibliographies are used as an additional entry point into the peer-reviewed journal literature, the ultimate repository of biological knowledge and the source of relevant annotations extracted from text and linked to the corresponding entries in biological databases (Kuruppu & Gruber, 2006).

Peer-Reviewed Journal Literature

Peer-reviewed journal articles are the principal means of scholarly communication in the biological sciences and an essential part of the scientific research process. Not only do researchers want to disseminate the results of their work to the public

and their peers but they also want to make sure their research findings are original. While the highlights of scientific discoveries are usually described in the mass media and later in books, the details and results of the most current research studies are largely reported through articles in peer-reviewed journals.

The importance of peer-reviewed journals as a means of scholarly communication is particularly evident in the biomedical sciences, including heavily research driven disciplines such as bioinformatics. Biomedical research is almost exclusively reported through articles in peer-reviewed journals for three reasons. The first is that research results are often time-sensitive. The expediency of access to current research findings can be crucial, both for scientists, who discover new drugs and treatments for these diseases, and for clinicians, who care for patients suffering from these conditions. The second is that biomedical research findings tend to go out of date more quickly than in other disciplines, largely because of the rapid changes in information technology and research methodology. Findings more than five years old are often considered to be out of date. Finally, peer-reviewed journals dominate the scholarly communication in the biomedical sciences because in academia, the publication of peer-reviewed articles, especially in the journals with high impact factors, is one of the most important criteria for promotion and tenure (Harley, Acord, Earl-Novell, Lawrence, & King, 2010).

Books

Although books are not a primary means for communicating research findings in bioinformatics, they continue to play an important role in supporting teaching and learning at the universities, especially textbooks at the undergraduate level. Additional publication types such as handbooks and lab manuals are used in relatively smaller numbers.

“Invisible College” Resources

Because the biological sciences is a “social, team-based enterprise,” an informal communication network of scholars, sometimes referred to as “invisible college,” is another important source of information for bioinformaticians (Harley, et al., 2010, p. 209). The concept of an invisible college originated in the 17th century in the Royal Society of London whose members did not belong to any formal institution and referred to themselves as an invisible college (Carey, 2011). It served as an influential channel for information exchange among scholars and allowed them to monitor scientific progress in their field, usually by the means of marginalia. Today, these channels include social media such as email, blogs, wikis, discussion lists, and society and professional hubs that have become widely accepted as platforms for scholarly communication. These informal methods of scholarly communication enable a wide range of interaction among scientists, such as allowing scholars to connect to each other and to discover new ideas, provide daily updates in the field, and facilitate discussions on research findings.

LIBRARY USE IN BIOINFORMATICS

In an attempt to understand what services libraries should provide to the bioinformatics community, surveys were conducted asking whether bioinformaticians and researchers in related bioscience fields, such as molecular biology and biomedicine, were aware of the resources in the library and whether they considered library collections as an important source of information (Brown, 2005; Grefsheim & Rankin, 2007; Haglund & Olsson, 2008; Hemminger, Lu, Vaughan, & Adams, 2007). These studies found that most bioscientists rarely utilized library collections as a primary source of information. The majority reported having learned about information sources, such as biological and journal literature databases, in their laboratories or

classrooms, from mentors and peers, rather than in campus libraries from librarians. Expertise and knowledge about these resources, and how to use them, were usually passed on by word of mouth to novice researchers by more experienced peers. While most participants had a high regard for journal literature as one of the primary sources of information, they did not fully utilize the library's bibliographic databases to find information. Many respondents located relevant journal articles through lists of journal references or through the PubMed's MEDLINE database, as well as by word of mouth, rather than through other, more 'library correct' search methods. Some researchers were either unaware of the library's databases or consulted these databases infrequently due to lack of time or search skills. Overall the participants made little use of traditional library collections. Of all the library collections used, full-text online journals were clearly preferred.

Lack of awareness of the valuable resources libraries have to offer or the skills in using these resources hinders the effective use of the university library by science scholars. A careful examination of the survey results suggests that in order to provide effective service to the bioinformatics user community, librarians should not only promote the value and usefulness of relevant library databases during instruction sessions, reference interactions, or research consultations but also discover effective methods for "collecting" freely available bioinformatics resources, such as molecular biology databases and data mining tools, and making them readily available through various venues that are already familiar to patrons, for example electronic journal lists or departmental websites.

BIOINFORMATICS COLLECTION DEVELOPMENT: RESOURCE TYPES, SELECTION CRITERIA, AND RECOMMENDATIONS

Core Values

Since the publication of *The Five Laws of Library Science* in 1931, the core values of librarianship have remained fundamentally the same: 1) Books are for use, 2) Every reader his or her book, 3) Every book its reader, 4) Save the time of the reader, and 5) The library is a growing organism (Ranganathan). Although in the digital era we can replace the term "book" with the term "information resource" and the term "reader" with the term "user," these core values continue to motivate our emphasis on the user and stress our need for a continuous readjustment, or "creative reimagining" (Horava, 2010, p. 151), of our ideas on what constitutes a useful library collection. It is critical to assimilate this knowledge into the collection building practices while maintaining the collection development core values—determining which resources best meet information needs of the library's users and adding the most suitable resources to the collection, whether these resources are physically located in a particular building or are made accessible by some other means. As the information needs of library users change and grow, so should a library collection. This need for continuous "creative reimagining" of a collection's content and scope is critical for a librarian charged with developing a collection in the rapidly evolving discipline of bioinformatics. Such a collection should be developed with strong awareness of how bioinformaticians seek information and communicate their research findings.

Characteristics of a Bioinformatics Library Collection

Some of the characteristics of a bioinformatics library collection are similar to those in other disciplines in the biological sciences:

1. The objective of the collection is to support teaching, learning, and research of the faculty and students of the university's programs in bioinformatics
2. Peer-reviewed journals are at the heart of a bioscience collection with other resource types being largely complementary
3. Expediency of access to the most current research resources is crucial in the biosciences; thus the emphasis should be put on providing access to electronic journals rather than on subscribing to print versions
4. Preference is given to journals with a high-impact factor, to journals included on a list of core journals for the discipline, and to journals requested by departmental faculty

The uniqueness of a bioinformatics collection lies in these two characteristics:

1. Its highly interdisciplinary nature, evolving at the interfaces among biology, mathematics, and the computer, biomedical, and environmental sciences
2. Its diversity, consolidating resources in various formats and from diverse sources, which, as a rule, are not considered as part of a traditional library collection but are essential in the field of bioinformatics

The specific subject focus of a collection is dependent on particular academic programs and research areas within an individual institution. Due to the interdisciplinary nature of the bioinformatics curriculum, acquisitions may need to be made in other subject areas, such as the computer and environmental sciences, statistics,

mathematics, and neurosciences. This may require co-ordination with other librarians. Subject area collection policies tailored to the university's academic programs can help outline the necessary standards for selecting, developing, and maintaining a discipline-specific collection.

Peer-Reviewed Journals

Bioinformatics research results and discoveries are primarily archived in biological and biomedical journals. In an academic setting, faculty sees the provision of peer-reviewed journals as one of the library's most valuable contributions to their teaching and research (Harley, et al., 2010). Online access to full-text peer-reviewed electronic journals is clearly preferred (Brown, 2005; Grefsheim & Rankin, 2007; Hemminger, et al., 2007; Kuruppu & Gruber, 2006). One of the main benefits of electronic journals, in the eyes of most bioinformaticians, is that these journals allow them to access published research results quickly from office or home (or wherever they happen to be working). Furthermore, electronic journals are usually available on the Web weeks before they are available in print, which is particularly important in the biological sciences where currency of content and expediency of access are often crucial. Other benefits of electronic journals include the following:

- a. support for distance education courses;
- b. searchability;
- c. hyperlinks to other publications;
- d. simultaneous use by multiple persons;
- e. suitability for adaptive technologies for visual impairment;
- f. personal/customized accounts.

Some scientists are also finding that electronic journals can offer significant value beyond full-text access by incorporating novel multimedia and interactive features in their content. For example, the *Journal of Visualized Experiments* (published

by JoVE) incorporates video based components (“video-articles”) in its content. Public Library of Science, a publisher of open access journals in the biomedical sciences, provides the ability for readers to rate and comment on, and to have discussion threads about, published articles. Though these innovations in some cases blur the lines between resource types, they represent a shift to greater reliance on digital media among scientists.

For libraries, some of the long-term benefits of electronic journals include reduction in shelving, processing, and binding costs, and, in some cases, savings on print subscription costs when a library is willing to cancel print subscriptions in favor of electronic access. Another benefit is increased availability and accuracy of usage statistics that helps with collection development decisions.

While most faculty and students in the biological sciences prefer electronic journals over their print counterparts, print journals are still a valuable component of a library collection. First, not all journals are available electronically. Some specialty areas may lag behind others in providing electronic access. Second, some library users still rely on print journals as a convenient means for browsing current issues of relevant journals in hard copy. Just as some readers prefer a print book over an electronic book, some still desire access to traditional print journals.

In some cases, maintaining selected titles in both print and electronic formats is justified. In the sciences, such titles include the journals *Nature* and *Science*, where the most important research across the biological sciences domain is being published. The availability of these flagship publications in their print version provides important benefits when computers are down. When both formats are not affordable, acquiring electronic journal titles is more critical in the biological sciences where currency and expediency of access is crucial.

Open Access Journals

The networked digital environment has enabled the creation of new types of resources that are successfully contributing to the scholarly communication system. A strengthening belief in the value of open access as an important practice of scholarly communication has taken firm hold in many disciplines. The rapid accumulation of knowledge and data in the biological sciences has resulted in an imperative to publish research results as quickly as possible (Harley, et al., 2010).

The open access publishing model is akin in spirit to the open source code movement in the computer programming field. Open access means that full-text research papers are available online as soon as they are published, free of charge, and where most restrictions on access or use have been lifted. Because of an effective online system for submission, peer reviewing, and publication that makes research papers rapidly available for a worldwide audience, the open access publishing practice has primarily gained acceptance in the biological sciences, where speed to publication is crucial, due to a persistent concern among researchers about the time lags between authorship and publication as well as about the high cost of the scientific journals, which prices many individual and institutional subscribers out of access to the most current research results. Free digital repositories, such as PubMed Central (<http://www.ncbi.nlm.nih.gov/pmc/>) at the National Library of Medicine, play a crucial role in ensuring the preservation of and access to published data.

Academic libraries play a critical role in fostering scholarly communication. They can further contribute to this process by incorporating open access journals into their collections. *BMC Bioinformatics* (BioMed Central) and *PLoS Computational Biology* (Public Library of Science)

are among the leading open access journals in the field and both publish more than 200 articles per year (Walters & Linnell, 2011). Bioinformatics librarians can collect open access journals by making them available through various venues that are familiar to patrons and that are likely to be used, for example through electronic journal lists and online catalogs (Collins & Walters, 2010). The links to open access journal packages can be added to the list of databases. This will ensure that faculty and students know how to access open access journals in their field quickly and conveniently. In turn, open access journals will show greater use “due to greater Web visibility and ease of discovery (and of course less cost)” (Mullen, 2010, p. 129).

Libraries can also be among those academic units that support the open access initiative through institutional memberships, subsidizing author publication fees for affiliated faculty, and thus encourage researchers to submit their work to open access journals. Adding open access journals to their collections can position libraries for leadership in such an important scholarly communication initiative and maximize the value of their collections.

Journal Selection Criteria

The general guidelines for developing most academic journal collections are also applicable to developing journal collection in bioinformatics:

1. Journals must be relevant to the information needs of library users
2. Journals must be scholarly (peer-reviewed or refereed)
3. Journals must clearly support teaching and research at the university
4. Preference is given to journals with high-impact factor, to journals included on a list of core journals for the discipline, and to journals requested by departmental faculty

Core Journals

The concept of a “core journal title” does not have universal applicability for all subject disciplines and all academic libraries. The purpose of an academic library is to collect resources that meet the needs of the user in that particular library. Thus, a core journal is a journal that is most relevant, or “indispensable” (Reitz, 2004, p. 183), for the users’ specialized research and learning needs, curriculum, as well as to the library’s mission and collecting priorities, which may differ from school to school.

Journal Impact Factors and Collection Development Decisions

Journal ranking provided by the *Journal Citation Reports (JCR)*, currently available through Thomson ISI’s *Web of Knowledge* (<http://apps.isiknowledge.com/>), plays an important evaluative role in identifying the journals that have been cited the most. In the sciences, journal impact factors are well known and considered important in author decisions about where to submit work as well as in the tenure review process. Some researchers, as well as some librarians, consider *JCR*’s ranked list of journals for a specific discipline to represent the core journal list for that discipline. The evaluative importance of citation count data is “predicated on the assumption that citation is evidence of contribution to scholarly communications and presumably an indicator of quality” (Nisonger, 2007, p. 57). Some suggest that the importance of journal impact factors is overrated (Monastersky, 2005). High-quality research papers also appear in journals with low impact factors as well as in journals with small circulation.

Faculty of 1000 (Faculty of 1000 Ltd.) (<http://f1000.com/>), an online post-publication literature awareness and peer review service, provides an alternative to the more conventional way of evaluating a journal’s importance in a particular

field. *Faculty of 1000* offers a rapid systematic review and rating of the most noteworthy primary research papers as judged by a panel of the world's leading researchers in the biological and medical sciences (called Faculty Members) and thus helps busy researchers cope with the abundance of published literature in the biological sciences. The service currently comprises 43 subject fields (called Faculties), including Bioinformatics & Computational Biology. One of the innovative features of *Faculty of 1000* is that it identifies and reviews important papers published not only in highly select journals such as *Nature*, *Science*, or *Cell*, but also in journals with small circulation, thus ensuring that papers are acknowledged for their research merit and not merely for the prestige of the journal in which they are published. The articles in less prestigious journals appear in the Faculty of 1000s' Hidden Jewels page (<http://www.f1000biology.com/top10/jewels/>). The key benefit of Faculty of 1000 for bioinformatics librarians is that it can help them identify the journals that publish key papers in the field.

Two journal titles, however, should be considered as the core titles for any academic collection in the biological sciences: *Nature* and *Science*. Apart from these two publications, the specialized research and teaching needs of a particular university should be the main guide for making journal selection decisions.

Books

A collection development policy, which guides book selection, a carefully crafted, coherent approval plan, and a subject profile policy can assist a bioinformatics librarian with book selection decisions to ensure that acquired titles fill the curriculum and research needs within the university. For some emerging fields within the bioinformatics discipline, such as neuroinformatics or ecoinformatics, where the body of knowledge is still in its infancy and the quantity of publication is still limited, it may be necessary to collect

all available material (if it otherwise meets the selection criteria).

The Library of Congress classification for bioinformatics and computational biology falls primarily within the QH range. Specific classifications include:

- QH426-QH470 (Genetics)
- QH506 (Molecular Biology)
- QH573-671 (Cell Biology)
- QP901-QP981 (Experimental Pharmacology)
- QR1-QR502 (Microbiology)
- QR355-QR502 (Virology)

Resources in bioinformatics may also be found in QA75-76 (Computer Science), QA1-939 (Mathematics), QD415-436 (Biochemistry), R (Medicine), and RC321-571 (Neurosciences).

Scholarly and professional works in the field should form the basis of the collection. Generally, lower undergraduate materials and textbooks are not purchased unless they offer a good overview treatment of a particular topic that is suited to all levels. Emphasis for collecting in bioinformatics is on current research, with the publication dates of most books purchased falling within the past three years. Acquisition of conference proceedings is important to the collection as it provides access to the most recent research in the field. Older materials are purchased if they are classics in bioinformatics or are requested by faculty. There is no specific geographical focus for the book collection. Although most materials collected are published in the United States, the content increasingly reflects global research in bioinformatics, especially in the developing countries.

Utilizing hard copies of print catalogs for book selection in bioinformatics has its advantages. They allow a librarian to mark or highlight selected titles as well as the catalog pages in place of book order forms. However, publishers often provide more information about their publications on their websites than they do in their print catalogs, including book reviews, tables of contents, text

excerpts, and links to new and forthcoming titles within the same subject area or topic. Elsevier, Wiley, American Society for Microbiology Press, Humana Press, Horizon Bioscience, Cold Spring Harbor Laboratory Press, and Springer publish high-quality books in bioinformatics and provide online catalogs on their Web sites as well as other helpful resources for librarians.

E-books have not yet gained a wide acceptance within the academic community (Zhang & Beckman, 2011). Among the primary concerns expressed by e-book users are difficulties in reading from the computer screen, problems related to locating e-books, and navigation issues. However, the value-added features of e-books are obvious: 24/7 desktop availability, concurrent access for multiple users, searching and saving features of e-books, downloadable graphics, capability to print and download sections or chapters, as well as the ability to bookmark, highlight, and annotate text. “Just as e-journals have found acceptance, e-books are likely to become more commonplace in the near future. The convenience factor alone will be a significant driver” (Bierman, Ortega, & Rupp-Serrano, 2010, p. 85). Among those patrons who can benefit from having access to e-books the most are off-site library users and distance education students.

Bibliographic Databases

A decision to subscribe to a bibliographic database for bioinformatics may involve several subject librarians. A collaborative team approach to the selection and evaluation of a database is important. Though many of the evaluation criteria with regard to the acquisition of an online bibliographic database focus on technological and licensing issues, the discussion of these issues is beyond the scope of this chapter. The primary role of a bioinformatics librarian in a database selection is evaluation of the quality of a database’s content

since the librarian brings to the selection decisions his or her knowledge of the specialized needs of bioinformatics faculty and students within the specific institution. This knowledge may lead to preferring one database platform over the other, given a choice.

The content of a database is evaluated by using the same criteria that govern the evaluation of other types of electronic library materials: relevance, authority, scope, coverage, and frequency of updates. A traditional database evaluation checklist is useful when evaluating the content of a database in bioinformatics. The following questions should be asked in order to make the best choices for the institution and users.

- Is a database’s publisher/vendor reputable?
- What are the dates of coverage?
- How often is the content updated?
- How unique is this database to the library’s current collection? Does it fill current gaps in the collection? Does it duplicate resources already provided by the library?
- If a database is available on more than one vendor platform, which platform will allow the user to locate information more quickly and efficiently?
- If the database is later canceled, will the library still have access to the licensed content?
- Does the database have a user-friendly search interface? Is it aesthetically attractive?

Preference is given to full-text databases that provide access to the most current scholarly content, including, when possible, access to pre-prints, and to those databases that enable such features as citation count, links to referenced publications, image representation, and convenient article printing options (such as a PDF file).

Molecular Biology Databases

Traditional library resources such as journals, books, and bibliographic databases are only a fraction of the current “information landscape” in bioinformatics where reliance on biological databases and data mining tools is prevalent (Geer, 2006, p. 286). There are a large number and diversity of publicly available biological databases and the number is growing, doubling in size about every 15 months (Cleveland & Cleveland, 2009, p. 240). Many are updated daily to accommodate newly submitted data and make it more readily accessible.

While these developments hold great value for research, bioinformaticians are challenged by the growing number and complexity of bioinformatics resources that seem “to outpace most researchers’ ability to keep up with the fast-moving field” (Geer, 2006, p. 296). Furthermore, identifying and locating new and emerging information resources is a time-consuming task for busy researchers (Cannata, Merelli, & Altman, 2005; Wren, 2004). Despite their self-sufficiency as information seekers, these researchers may benefit from the libraries’ traditional role as information managers who can help organize and facilitate access to the bioinformatics “resourceome” (Cannata, et al., 2005, e76).

Given the number and diversity of bioinformatics resources, the relative difficulty of locating these resources, the scarcity of available location guides, and limitations in annotation, academic libraries are beginning to recognize the need to identify, organize, annotate, and better facilitate access to these resources in order to support the bioinformatics programs at their institutions (Chen, Chattopadhyay, Bergen, Gadd, & Tannery, 2007; Minie, et al., 2006). Few academic libraries have well-developed specialized research portals that act as information gateways to selected, organized, and annotated bioinformatics resources. Among these libraries are the Health Sciences Library System (HSLS) at the University of Pittsburgh,

who have developed the Online Bioinformatics Resources Collection (OBRC) (<http://www.hslls.pitt.edu/guides/genetics/obrc>), and the University Libraries of George Mason University, who have created the Bioinformatics Research Portal (<http://gmutant.gmu.edu/melange/>). These portals, combining internal and external bioinformatics resources, can be viewed as the natural extension of the library’s collection in bioinformatics as well as the expansion of roles for librarians looking for new ways to contribute to this emerging field.

There are many authoritative lists of bioinformatics resources that can assist a librarian with identifying and selection of biological databases and tools (Benson, 2011; Brooksbank, Cameron, & Thorton, 2010; Galperin & Cochrane, 2011; Sayers, et al., 2011). The journal *Nucleic Acids Research*, published by Oxford University Press, devotes the first issue of each year (the Database Issue) to an in-depth description of new and enhanced biological databases. The Database Issue also includes a comprehensive list of databases called Database Collection, available at <http://www.oxfordjournals.org/nar/database/a/>. In its January 2011 issue the Database Collection listed 1330 biological databases.

EDUCATIONAL OPPORTUNITIES FOR BIOINFORMATICS LIBRARIANS

Collection development for the complex and rapidly evolving field of bioinformatics can be a challenge. Bioinformatics librarians must not only be aware of a broad range of available bioinformatics resources, including biological databases and data mining tools, but also possess at least some degree of subject knowledge obtained either through formal education or self-education. This challenge provides the opportunity for a librarian who is looking for ways to contribute to research efforts across the discipline by building a useful collection.

Kajosalo (2006) identified three levels of a bioinformatics librarian's expertise required of a librarian to provide a bioinformatics user with the best assistance possible: basic, intermediate, and expert. Although the intermediate and expert levels of service, according to Kajosalo, would require a librarian to have a strong subject matter proficiency, particularly in the areas of providing in-depth reference service, consulting, and specialized instruction, a librarian at the basic level of service should already have good knowledge of bioinformatics databases, be able to expand print and online library collections, and maintain a curated portal of bioinformatics online resources.

While it is easier for a librarian with a background in the biological and/or computer sciences to advance in the area of bioinformatics collection development, various training programs, and workshops are becoming more readily available for information professionals from any background. The NCBI's Educational Resources webpage of the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/education/>) and the Medical Library Association's Continuing Education Clearinghouse (<http://cech.mlanet.org/>) offer additional information about training and education opportunities for librarians and other professionals. For nearly two decades, the Marine Biological Laboratory (MBL) hosts the BioMedical Informatics Course (http://www.mbl.edu/education/courses/special_topics/med.html). This course, sponsored by the National Library of Medicine, is held twice a year at the MBL and is committed to providing continuing education for bioinformaticians and other scholars in informatics.

For those seeking formal academic education, the International Society for Computational Biology (ISCB) (<http://www.iscb.org/iscb-degree-certificate-programs>) provides information about major bioinformatics programs around the world. Some library science schools either offer a separate master's degree in bioinformatics, for example, the Graduate School of Library and Information

Science (GSLIS) at the University of Illinois (<http://www.lis.illinois.edu/academics/programs/ms-bioinformatics>) or provide bioinformatics training as part of their programs, for example the School of Information and Library Science (SILS) at the University of North Carolina (UNC) at Chapel Hill, which awards a certificate of specialization in bioinformatics in conjunction with either a master's of library science or a master's of information science (<http://sils.unc.edu/programs/certificates/bioinformatics>).

While training and/or formal education can certainly enhance a librarian's skills, most skills are learned on the job. Mastery comes through frequent, routine, and repeated collection development activities. Familiarity with selection tools, techniques, and resources are, of course, essential for success.

The following set of skills and competencies will equip a librarian to meet the challenge of developing library collections in bioinformatics:

- Familiarity with the curricula within the academic department and the research activities of the departmental faculty
- At least some degree of subject knowledge obtained either through training, formal education, or self-education
- Awareness of current research trends and discoveries in the field
- Knowledge of the publishing industry in the biological sciences
- Familiarity with scholarly communication trends
- Knowledge of selection tools and selection criteria in the life sciences
- Knowledge of automated collection development tools
- Time and budget management skills
- Negotiation and communication skills

The complex matter of intellectual property management may also be part of the new knowledgebase expected of academic librarians in any

subject area as well as the knowledge of Web design and the ability to produce blogs, wikis, websites, and other electronic methods for communicating with library users. Staying on top of these expectations is challenging but essential. Such personal traits as flexibility, open-mindedness, ability to embrace change, and commitment to continual learning must be layered inside professional competence in order to effectively develop library collections in the networked era.

CONCLUSION

It is becoming increasingly difficult to define the actual borders of library collections in the hybrid print/electronic/open access environment as access gradually overtakes ownership as the model for many academic libraries. Furthermore, each academic research setting is unique, and bioinformatics collections across libraries can be similarly diverse. Horava (2010) observes that today's collection is "everywhere and nowhere—it is a cloud of distributed resources in a variety of places around the globe that are made centrally available via the library" (Horava, 2010, p. 151). Yet, even today it is possible to recognize Ranganathan's vision of the library as a growing organism as the main incentive that drives librarians to proactively stay abreast of evolving information environment while maintaining the library's core values.

ACKNOWLEDGMENT

I would like to acknowledge the support of the University Libraries at George Mason University who has granted me research leave to work on this chapter. I would also like to acknowledge the assistance and encouragement of my husband, James Martin, who provided valuable comments and suggestions as I shepherded this chapter through its development.

REFERENCES

- Altman, R. B., & Mooney, S. D. (2006). Bioinformatics. In Shortliffe, E., & Cimino, J. (Eds.), *Biomedical Informatics: Computer Applications in Health Care and Biomedicine* (pp. 762–789). New York, NY: Springer.
- Benson, G. (2011). Web server issue. *Nucleic Acids Research*, 39(2), W1–W2. doi:10.1093/nar/gkr515
- Bierman, J., Ortega, L., & Rupp-Serrano, K. (2010). E-book usage in pure and applied sciences. *Science & Technology Libraries*, 29(1-2), 69–91. doi:10.1080/01942620903579393
- Brooksbank, C., Cameron, G., & Thornton, J. (2010). The European bioinformatics institute's data resources. *Nucleic Acids Research*, 38, D17–D25. doi:10.1093/nar/gkp986
- Brown, C. (2005). Where do molecular biology graduate students find information? *Science & Technology Libraries*, 25(3), 89–104. doi:10.1300/J122v25n03_06
- Buehler, L., & Rashidi, H. (Eds.). (2005). *Bioinformatics basics: Applications in biological science and medicine*. Boca Raton, FL: Taylor & Francis.
- Cannata, N., Merelli, E., & Altman, R. B. (2005). Time to organize the bioinformatics resourceome. *PLoS Computational Biology*, 1, e76. doi:10.1371/journal.pcbi.0010076
- Carey, J. (2011). Faculty of 1000 and VIVO: Invisible colleges and team science. *Issues in Science & Technology Librarianship*, 65. Retrieved from <http://www.istl.org/11-spring/article1.html>.
- Chen, Y. B., Chattopadhyay, A., Bergen, P., Gadd, C., & Tannery, N. (2007). The online bioinformatics resources collection at the University of Pittsburgh health sciences library system - A one-stop gateway to online bioinformatics databases and software tools. *Nucleic Acids Research*, 35(1), D780–D785. doi:10.1093/nar/gkl781

- Chiang, K. S. (2004). Biology databases for the new life sciences. *Science & Technology Libraries*, 25(1/2), 139–170. doi:10.1300/J122v25n01_10
- Cleveland, A. D., & Cleveland, D. B. (2009). *Health informatics for medical librarians*. New York, NY: Neal-Schuman.
- Collins, C. S., & Walters, W. H. (2010). Open access journals in college library collections. *The Serials Librarian*, 59(2), 194–214. doi:10.1080/03615261003623187
- Cummings, M., & Temple, G. (2010). Broader incorporation of bioinformatics in education: Opportunities and challenges. *Briefings in Bioinformatics*, 11(6), 537–554. doi:10.1093/bib/bbq058
- Dennie, D. (2010). The provision of bioinformatics services in Canadian academic libraries. *Journal of the Canadian Health Libraries Association*, 31(3), 99–107. doi:10.5596/c10-028
- Dinkelman, A. L. (2007). See a need, fill a need - Reaching out to the bioinformatics research community at Iowa State University. *Issues in Science & Technology Librarianship*, 52. Retrieved from <http://www.istl.org/07-fall/refereed1.html>.
- Fenstermacher, D. (2005). Introduction to bioinformatics. *Journal of the American Society for Information Science and Technology*, 56(5), 440–446. doi:10.1002/asi.20133
- Fulekar, M. (Ed.). (2009). *Bioinformatics: Applications in life and environmental sciences*. New York, NY: Springer. doi:10.1007/978-1-4020-8880-3
- Furge, L. L., Stevens-Truss, R., Moore, D. B., & Langeland, J. A. (2009). Vertical and horizontal integration of bioinformatics education. *Biochemistry and Molecular Biology Education*, 37(1), 26–36. doi:10.1002/bmb.20249
- Galperin, M. Y., & Cochrane, G. R. (2011). The 2011 nucleic acids research database issue and the online molecular biology database collection. *Nucleic Acids Research*, 39(1), D1–D6. doi:10.1093/nar/gkq1243
- Geer, R. (2006). Broad issues to consider for library involvement in bioinformatics. *Journal of the Medical Library Association*, 94(3), 286–298.
- Grefsheim, S. F., & Rankin, J. A. (2007). Information needs and information seeking in a biomedical research setting: A study of scientists and science administrators. *Journal of the Medical Library Association*, 95(4), 426–434. doi:10.3163/1536-5050.95.4.426
- Haglund, L., & Olsson, P. (2008). The impact on university libraries on changes in information behavior among academic researchers: A multiple case study. *Journal of Academic Librarianship*, 34(1), 52–59. doi:10.1016/j.acalib.2007.11.010
- Hancock, J., & Zvelebil, M. (Eds.). (2004). *Dictionary of bioinformatics and computational biology*. Hoboken, NJ: Wiley-Liss. doi:10.1002/0471650129
- Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S., & King, C. (2010). *Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines*. Retrieved from <http://escholarship.org/uc/item/15x7385g;jsessionid=4CF15BA02B93D9A10C0E2D5F69D2D20F>.
- Hemminger, B., Losi, T., & Bauers, A. (2005). Survey of bioinformatics programs in the United States. *Journal of the American Society for Information Science and Technology*, 56(5), 529–537. doi:10.1002/asi.20141
- Hemminger, B. M., Lu, D., Vaughan, K. T. L., & Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 58(14), 2205–2225. doi:10.1002/asi.20686

- Horava, T. (2010). Challenges and possibilities for collection management in a digital age. *Library Resources & Technical Services*, 54(3), 142–152.
- Kajosalo, E. (2006). *Bioinformatics: An instructional opportunity for academic science and engineering libraries*. Paper presented at the American Chemical Society National Meeting. San Francisco, CA. Retrieved from <http://acscinf.org/docs/meetings/232nm/presentations/232nm35.pdf>.
- Kawano, S., Ono, H., Takagi, T., & Bono, H. (2011, July 29). Tutorial videos of bioinformatics resources: Online distribution trial in Japan named TogoTV. *Briefings in Bioinformatics*.
- Kuruppu, P., & Gruber, A. (2006). Understanding the information needs of academic scholars in agricultural and biological sciences. *Journal of Academic Librarianship*, 32(6), 609–623. doi:10.1016/j.acalib.2006.08.001
- MacMullen, W., & Denn, S. (2005). Information problems in molecular biology and bioinformatics. *Journal of the American Society for Information Science and Technology*, 56(5), 447–456. doi:10.1002/asi.20134
- Minie, M., Bowers, S., Tarczy-Hornoch, P., Roberts, E., James, R., Rambo, N., & Fuller, S. (2006). The University of Washington health sciences library biocommons: An evolving north-west biomedical research information support infrastructure. *Journal of the Medical Library Association*, 94(3), 321–329.
- Monastersky, R. (2005). The number that's devouring science. *The Chronicle of Higher Education*, 52(8). Retrieved from <http://chronicle.com/free/v52/i08/08a01201.htm>
- Mullen, L. B. (2010). *Open access and its practical impact on the work of academic librarians: Collection development, public services, and the library and information science literature*. Oxford, UK: Chandos. doi:10.1533/9781780630229
- Nisonger, T. (2007). Journals in the core collection. *The Serials Librarian*, 51(3/4), 51–73. doi:10.1300/J123v51n03_05
- Peterson's. (2011). *Peterson's graduate programs in the biological sciences*. Princeton, NJ: Peterson's.
- Ranganathan, S. (2005). Bioinformatics education - Perspectives and challenges. *PLoS Computational Biology*, 1(6), 447–448. doi:10.1371/journal.pcbi.0010052
- Ranganathan, S. R. (1931). *The five laws of library science*. London, UK: Edward Goldston.
- Rapp, B., & Wheeler, D. (2005). Bioinformatics resources from the national center for biotechnology information: An integrated foundation for discovery. *Journal of the American Society for Information Science and Technology*, 56(5), 538–550. doi:10.1002/asi.20142
- Rédei, G. P. (2008). *Encyclopedia of genetics, genomics, proteomics, and informatics*. New York, NY: Springer.
- Reitz, J. M. (2004). *Dictionary for library and information science*. Westport, CT: Libraries Unlimited.
- Saravanana, V., & Shanmughavel, P. (2007). E-learning as a new tool in bioinformatics teaching. *Bioinformation*, 2, 83–85. doi:10.6026/97320630002083
- Sayers, E., Barrett, T., Benson, D., Bolton, E., Bryant, S., & Canese, K. (2011). Database resources of the national center for biotechnology information. *Nucleic Acids Research*, 39(1).
- Schneider, M., Watson, J., Attwood, T., Rother, K., Budd, A., & McDowall, A. (2010). Bioinformatics training: A review of challenges, actions and support requirements. *Briefings in Bioinformatics*, 11(6), 544–551. doi:10.1093/bib/bbq021

Tennant, M. (2005). Bioinformatics librarian: Meeting the information needs of genetics and bioinformatics researchers. *RSR. Reference Services Review*, 33(1), 12–19. doi:10.1108/00907320410519333

Tramontano, A. (2009). *Bioinformatics*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0001900.pub2/full>.

Vaidhyanathan, V. (2008). *Why can't it all be on the web? The information needs of biomedical informatics scientists*. Unpublished Master's Thesis. Chapel Hill, NC: University of North Carolina.

Walters, W. H., & Linvill, A. C. (2011). Characteristics of open access journals in six subject areas. *College & Research Libraries*, 72, 372–392.

Weisman, D. (2010). Incorporating a collaborative web-based virtual laboratory in an undergraduate bioinformatics course. *Biochemistry and Molecular Biology Education*, 38(1), 4–9. doi:10.1002/bmb.20368

Wren, J. D. (2004). 404 not found: The stability and persistence of URLs published in MEDLINE. *Bioinformatics (Oxford, England)*, 20, 668–672. doi:10.1093/bioinformatics/btg465

Zatz, M. (2002). Bioinformatics training in the USA. *Briefings in Bioinformatics*, 3, 353–360. doi:10.1093/bib/3.4.353

Zhang, D., Zhao, L., & Nunamaker, L. (2004). Can e-learning replace classroom learning? *Communications of the ACM*, 47, 74–79. doi:10.1145/986213.986216

Zhang, Y., & Beckman, R. (2011). E-book usage among chemists, biochemists and biologists: Findings of a survey and interviews. *Issues in Science & Technology Librarianship*, 65. Retrieved from <http://www.istl.org/11-spring/article2.html>.

KEY TERMS AND DEFINITIONS

Bioinformatics: The application of computers and computational techniques for the interpretation of biological data.

DNA: Deoxyribonucleic acid, the hereditary material in most organisms, including humans.

e-Learning: The use of electronic media and Internet technologies to enhance educational activities.

Genome: An organism's complete hereditary information, including all of its genes.

APPENDIX: KEY RESOURCES

Journals

- **Bioinformatics** (Oxford University Press) <http://bioinformatics.oxfordjournals.org/>. The leading journal in the field publishing research papers and review articles that focus primarily on new developments in genome bioinformatics and computational biology.
- **BMC Bioinformatics** (BioMed Central) <http://www.biomedcentral.com/bmcbioinformatics>. An open access, peer-reviewed journal publishing research articles on all aspects of bioinformatics and computational biology.
- **Briefings in Bioinformatics** (Oxford University Press) <http://bib.oxfordjournals.org/>. Publishes reviews of biological databases and analytical tools for genetics and molecular biology. Provides user help and guidance to the non-specialist.
- **Database: The Journal of Biological Databases and Curation** (Oxford University Press) <http://database.oxfordjournals.org/>. An online-only open access journal publishing descriptions of databases, database tools, and significant updates to established databases, as well as user tutorials and technical notes on database development.
- **Journal of Bioinformatics and Computational Biology (JBCB)** (World Scientific) <http://www.worldscinet.com/jbcb/>. Publishes research, expository tutorial, and review papers covering the analysis of cellular information.
- **PLoS Computational Biology** (Public Library of Science). <http://www.ploscompbiol.org>. An open-access, peer-reviewed journal publishing research papers “of exceptional significance” on all aspects of research in computational biology.

Bibliographic Databases

- **BIOSIS** (Thomson Reuters). The world’s largest database of life science and biomedical research literature, including journals, conferences, patents, books, and review articles. Includes over 12 million bibliographic records dating back to 1926. Most of the records include abstracts written by the author.
- **Inspec** (Institution of Engineering and Technology [IET]). Includes citations and abstracts from publications in the fields of physics, engineering, computer science, information technology, nanotechnology, biomedical technology, and biophysics that may be of particular interest to software and database developers.
- **MEDLINE** (National Library of Medicine). The National Library of Medicine’s premier bibliographic database (often cited as PubMed). Contains over 21 million references to journal articles in biomedical sciences. Available either through PubMed (free of charge) or from platforms provided by different vendors, including Ebsco, Ovid, and Cambridge Scientific Abstracts.

Protocols

- **Current Protocols in Bioinformatics** (Wiley) <http://www.currentprotocols.com/bioinformatics>. Provides information on the types of data and protocols required to perform basic tasks in the area of bioinformatics. Offers step-by-step instructions for working with applications and databases as well as guidance on selecting software parameters, analyzing data, interpreting results, and generating hypotheses. Includes glossary of bioinformatics terms.

Collections of Molecular Biology Databases and Bioinformatics Tools

- **Bioinformatics Links Directory:** http://bioinformatics.ca/links_directory/. An open access repository of curated bioinformatics links selected on the basis of recommendations from the experts in the field and published in collaboration with *Nucleic Acids Research* (Oxford University Press).
- **Database Collection:** <http://www.oxfordjournals.org/nar/database/a/>. The *Nucleic Acids Research* (Oxford University Press) comprehensive collection of new and enhanced molecular biology databases described in the journal's Database Issue, published in January of each year, as well as a selection of other molecular biology databases that are freely available on the Web and might be useful to the molecular biologist.
- **The European Bioinformatics Institute (EBI) resources:** <http://www.ebi.ac.uk/>. One of the world's most important collections of biological data, including DNA sequences (EMBL-Bank) (<http://www.ebi.ac.uk/embl/>), protein sequences (UniProt) (<http://www.ebi.ac.uk/uniprot/>), animal genomes (Ensembl) (<http://useast.ensembl.org/index.html>), three-dimensional structures (the Protein Data Bank in Europe) (<http://www.ebi.ac.uk/pdbe/>), data from gene expression experiments (ArrayExpress) (<http://www.ebi.ac.uk/arrayexpress/>), and protein-protein interactions (IntAct) (<http://www.ebi.ac.uk/intact/>).
- **ExPASy: Bioinformatics Resource Portal:** <http://www.expasy.ch/>. Created and maintained by the Swiss Institute of Bioinformatics (SIB), provides access to more than 120 SIB bioinformatics resources, including a protein database UniProtKB/Swiss-Prot, a human-centric protein knowledge platform neXtProt, and a database of protein interactions STRING.
- **The National Center for Biotechnology Information (NCBI) resources:** <http://www.ncbi.nlm.nih.gov>. A collection of 39 literature and molecular biology databases, among which are GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), containing nucleotide sequences, the related database RefSeq (<http://www.ncbi.nlm.nih.gov/RefSeq/>), containing sequences and their protein products, OMIM (Online Mendelian Inheritance in Man) (<http://www.ncbi.nlm.nih.gov/omim>), a catalog of human genetic disorders and their causative genes, PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>), which comprises more than 21 million citations for biomedical literature from the database MEDLINE produced by the National Library of Medicine, PubMed Central (PMC) (<http://www.ncbi.nlm.nih.gov/pmc/>), a free digital archive of full-text life sciences journal articles, the NCBI Bookshelf, a collection of biomedical books adapted for the Web (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=Books>), and an array of educational materials and resources (<http://www.ncbi.nlm.nih.gov/Education/>).

- **Pathguide:** <http://www.pathguide.org/>. Provides an overview of more than 300 biological pathway and network databases, including databases on metabolic pathways, signaling pathways, transcription factor targets, gene regulatory networks, genetic interactions, protein–compound interactions, and protein–protein interactions.

Bioinformatics Portals in Academic Libraries

- **Bioinformatics Research Portal** (George Mason University): <http://gmutant.gmu.edu/melange/>. Combines the features of a blog, a resource database, a search engine, a collection of useful links, and RSS feeds for the tables of contents from the latest issues of selected bioinformatics journals. Integrates library-owned materials with free Internet resources and provides access to both subscribed and free content in one convenient place.
- **Online Bioinformatics Resources Collection (OBRC)** (University of Pittsburgh): <http://www.hslls.pitt.edu/obrc/>. Includes online bioinformatics databases and software tools, most of which have been published by *Nucleic Acids Research* or listed in its Database Collection, as well as selected databases and software tools described in other journals, such as *Bioinformatics* and *BMC Bioinformatics*, and a number of unpublished online software tool.

Educational Resources

- **BioMedical Informatics Course** (Marine Biological Laboratory): http://www.mbl.edu/education/courses/special_topics/med.html. A week-long course sponsored by the National Library of Medicine. Directed at librarians, clinicians, educators, and administrators.
- **MLA Educational Clearinghouse** (Medical Library Association): <http://cech.mlanet.org/>. The Medical Library Association's listing of courses, educational activities, and educational resources for practicing biomedical librarians.
- **The NCBI Bookshelf:** <http://www.ncbi.nlm.nih.gov/sites/entrez?db=Books>. A collection of biomedical books, many of which are on-line versions of standard biomedical textbooks on cell biology, molecular biology and biochemistry. Increasingly serves as the repository for the NCBI documentation, tutorials, course materials, and newsletters.
- **The NCBI Education page:** <http://www.ncbi.nlm.nih.gov/Education/>. The primary portal for accessing educational materials and resources from the National Center for Biotechnology Information (NCBI). Provides Web, print and video tutorials, interactive tools and an archive of course material. Includes materials designed for new users, such as About NCBI and the NCBI Guide. Produces the online NCBI News and offers RSS feeds and mailing lists, along with a presence on Facebook, Twitter and YouTube.
- **OpenHelix:** <http://www.openhelix.com/>. Provides a collection of various bioinformatics training materials, including narrated online tutorials, videos, and animated PowerPoint slides.