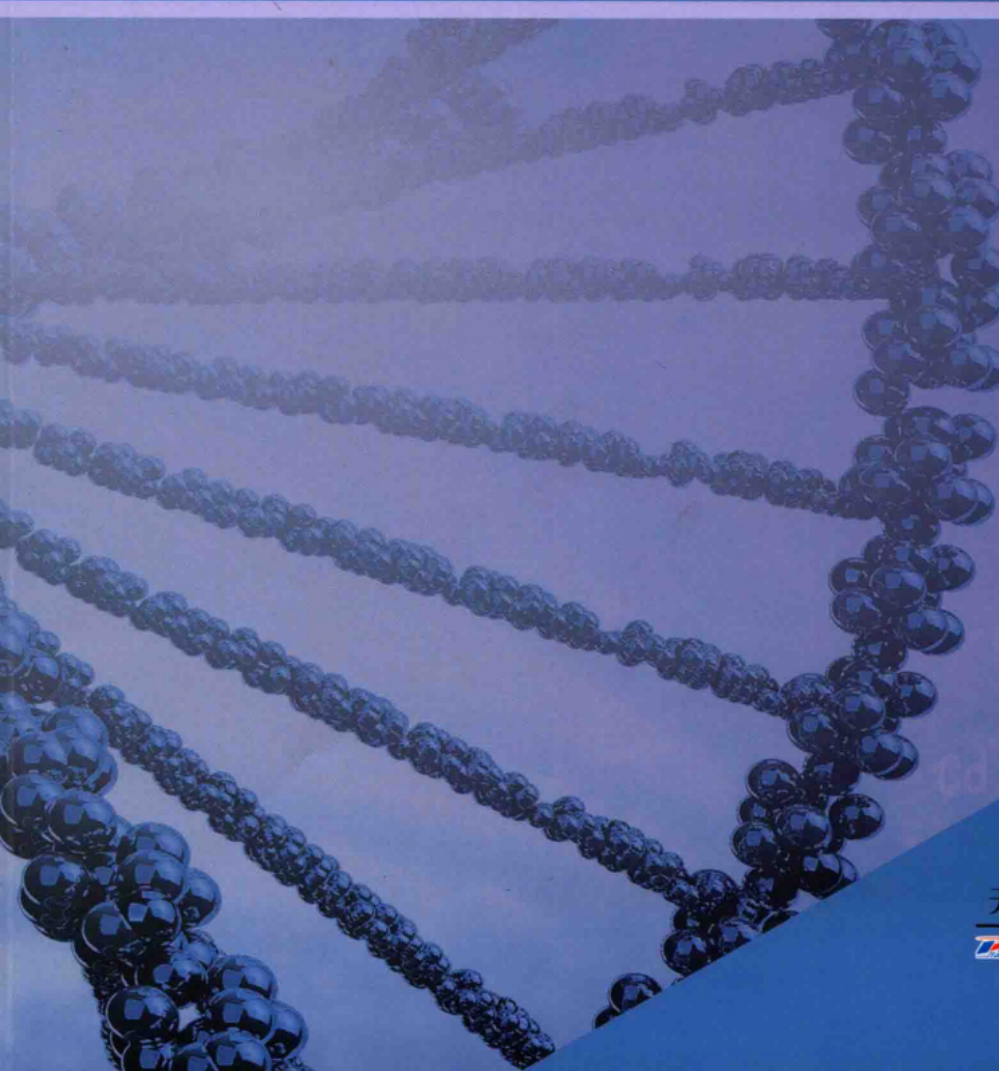




网络生物医学信息 理论与实践

主编 ● 朱金玲



天津出版传媒集团

 天津科学技术出版社



网络生物医学信息 理论与实践

主编◎朱金玲

天津出版传媒集团

 天津科学技术出版社

图书在版编目(CIP)数据

网络生物信息理论与实践 / 朱金玲主编. — 天津 :

天津科学技术出版社, 2017.1

ISBN 978-7-5576-2131-5

I. ①网… II. ①朱… III. ①生物信息论-应用-计
算机网络-网络安全-网络理论 IV. ①TP393.08

中国版本图书馆 CIP 数据核字(2017) 第 003019 号

责任编辑:石 崑

责任印制:兰 毅

天津出版传媒集团

 天津科学技术出版社

出版人:蔡 颢

天津市西康路 35 号 邮编 300051

电话(022)23332390

网址:www.tjkjcs.com.cn

新华书店经销

三河市腾飞印务有限公司印刷

开本 787×1092 1/16 印张 18.5 字数 400 000

2017 年 1 月第 1 版第 1 次印刷

定价:49.60 元

《网络生物医学信息理论与实践》编委会

主 编

朱金玲

副主编

张玉萍 吕冬霞

主 审

王景涛

编委名单

(按姓氏笔画排序)

王长山 王景涛 朱金玲

吕冬霞 张玉萍 张 虎

张宝成 金岳雷 侯 霞

主编简介

朱金玲,女,1967年出生。1990年毕业于东北师范大学生物系,获学士学位。同年就业于佳木斯大学,2004年获得佳木斯大学遗传学专业硕士学位。

现任基础医学院生物化学与细胞分子生物学部主任,教授,硕士研究生导师,《中国组织工程研究与临床康复》杂志编委,中国细胞生物学会医学细胞生物学分会委员。主要研究方向为医学生物学。现在佳木斯大学基础医学院从事细胞生物学、医学遗传学、生物信息学和分子生物学技术等本科生、研究生和留学生的教学和科研工作,本人是省级精品课《细胞生物学》的主讲教师,校级精品课《医学遗传学》的主持人。

近些年获得各种荣誉多项:如优秀教师、师德先进个人、十佳教学等荣誉称号。

近些年公开发表教科研论文30余篇,参与编写教材及专著8部。2012年参加国家自然科学基金1项,第3名,2015年参加1项,第四名。2007年、2013年主持省自然科学基金2项,《脂肪肝大鼠模型的建立及L-FABP在鼠肝细胞中的动态表达》;《体外细胞转化电化学试验方法的研究》,参加省自然科学基金等课题的研究4项。先后5次主持和参与获得省高校科学技术进步奖、省卫生厅科技奖、佳木斯市科技进步奖、校级科技进步奖。

前 言

生物医学信息学,按其字面意思,可以分为生物信息学和医学信息学。生物信息学是研究生物信息的采集、处理、存储、传播、分析和解释等各方面的一门学科,它通过综合利用生物学,计算机科学和信息技术揭示大量而复杂的生物数据所赋有的生物学奥秘。它是当今生命科学和自然科学的重大前沿领域之一,同时也将是 21 世纪自然科学的核心领域之一。生物信息学研究重点主要体现在基因组学和蛋白质组学两方面,具体说就是从核酸和蛋白质序列出发,分析序列中表达的结构功能的生物信息。

医学信息学,是医学和计算机科学的结合科学,是医学发展的必经阶段,也是一门新兴交叉学科。在信息学中,大致可以分出三个不同的研究层次:基础计算机科学、应用方法信息学以及应用信息学。医学信息学其研究内容在实际临床也有许多应用。比如,电子病历,这个系统可支持使用者获得完整、准确的资料;提示和警示医疗人员;给予临床决策服务;连接管理、书刊目录、临床基础知识以及其他设备。另外,还有医学影像的分析与鉴定,远程医疗等等。随着计算机的普及,这也更加完善了医学信息学的功能。

生物医学信息学是一门范围宽泛的学科,致力于在生物学研究、生物医学科学、医学以及医疗保健领域的工作实践当中,对于计算机科学、信息科学、信息学、认知科学以及人机交互的研究与应用。其他的许多领域,包括生物信息学、转化研究信息学、临床研究信息学、临床信息学、公共卫生信息学以及医学信息学在内,通常都属于是生物医学信息学的兄弟领域或者生物医学信息学之中的子领域。

信息化是当今世界发展的潮流和特征,正如世界未来学家阿尔文·托夫勒(Alvin Toffler)预言“谁掌握了信息、控制了网络,谁就掌握了整个世界。”如今,进入互联网已很简单,只要一台电脑、一个调制解调器和一根电话线,向地区互联网服务中心申请就可加入互联网络。向互联网索取的信息越来越多了。Internet 席卷全球的成功,证明了信息资源共享、信息服务网路化已经是不可逆转的潮流。网络环境下实现了信息的分布式存取,使得人们不受地域和时间的限制对全球范围内的信息资源进行查找和利用,因此获取了利用信息资源的能力显得越来越重要。网络环境下的现代生物医学的研究要求提供相应的网络化的信息服务,为此我们编写了《网络生物医学信息理论与实践》这本书。本书由佳木斯大学2015年教材建设规划项目资助出版。本书总共9章,以医学生物信息数据库为中心,主要介绍了国内外的知名度高、可信度强、应用范围广的生物医学数据库或网站,同时简要介绍了各种数据库的使用方法。

本书可作为医学、药学、医学信息学、生物医学工程和公共卫生等相关专业本科、专科的生物医学信息学课程教材或参考书,对健康信息学感兴趣的医生、护士、管理人员,以及从事医疗服务的技术人员、医疗保险事业人士也可以参考阅读。

通过阅读本书,可掌握生物医学信息学基本原理,能用生物医学信息学的方法和计算机技术开展医学研究。《网络生物医学信息学理论与实践》可作为生物学、医学、药学、医学信息学、生物医学工程和公共卫生等相关专业本科、专科的生物医学信息学课程教材或参考书,对健康信息学感兴趣的医生、护士、管理人员,以及从事生物和医学研究的人员具有指导意义。

朱金玲

2015年12月

目 录

第一章 网络生物医学信息资源概论

第一节 网络生物医学资源的类别、特点和评价	001
一、网络生物医学信息资源的类别	001
二、网络生物医学信息资源的特点	005
三、网络生物医学信息资源的评价	007
第二节 文献信息检索语言	011
一、文献检索语言的涵义、组成和类型	012
二、检索语言在网络信息组织与检索中的应用	016
第三节 一般检索技术与策略	018
一、一般检索技术	018

第二章 搜索引擎

第一节 搜索引擎概述	023
一、收索引擎的概念和发展现状	023
二、收索引擎的原理及分类	025
第二节 普通搜索引擎	027
一、中文普通搜索引擎	027
二、英文普通搜索引擎	031
第三节 生物医学专业搜索引擎	035

一、中国生物医学专业搜索引擎	035
二、国外生物医学专业搜索引擎	037
第四节 搜索引擎使用的方法和技巧	041
一、简单查询	041
二、双引号使用	041
三、搜索逻辑命令	042

第三章 网络文献类型

第一节 网络电子图书	044
一、网络电子图书出版	044
二、网络电子图书资源与数字图书馆	044
第二节 网络学术期刊	045
一、国外生物信息学术期刊	045
二、网络外文生物信息全文期刊	046
第三节 专利信息资源	049
一、中国专利	049
二、日本专利	050
三、美国专利	050
四、欧洲专利	051
第四节 网络学术会议与学位论文资源	051
一、网络学术会议	051
二、学位论文资源	053
第五节 网络医学多媒体资源	054
一、解剖图谱资源——VHDS	054
二、医学影像资源	055
三、心胸部影像数据库	055
四、生物医学图像处理精品课程	055
五、医学多媒体资源网	055

第四章 著名外文生物医学文献信息资源与检索系统

第一节 MEDLINE/PubMed	057
一、MEDLINE 文献收录和评价原则	057
二、MEDLINE 数据库特点	058
三、MEDLINE 的检索途径	058
四、Pubmed 检索方法与技巧	060
第二节 MEDLINEplus	063
一、Medline Plus 特点	064
二、MedlinePlus 资源	064
三、MedlinePlus 与 Medline/PebMed 的比较	067
第三节 荷兰《医学文摘》(EMBASE)	067
一、EMBASE.com 及其功能	067
二、EMBASE 主题词表	068
三、EMBASE.com 检索途径	068
四、检索结果处理	072
五、EMBASE 与 MEDLINE 的区别	073
第四节 BIOSIS Previews 网络数据库	074
一、BIOSIS Previews 数据库概况	074
二、BIOSIS Previews 数据库检索与链接	074
三、Biosis Previews 与 MEDLINE、EMBASE 比较	079
第五节 化学文摘网络版 SciFinder Scholar	079
一、CA 网络版 SciFinder Scholar 的特点	080
二、CA 网络版 SciFinder Scholar 的特色服务	082
三、CA 网络版 SciFinder Scholar 检索方法和使用技巧	082
第六节 OVID 数据库系统	088
一、Ovid 医学信息平台的特点	089
二、OvidSP 检索途径及方法	089

第七节 ISI Web of Knowledge 平台数据库	093
一、ISI Web of Knowledge 平台概述	093
二、ISI Web of Knowledge 平台检索技术和方式	095
三、Web of Science	097
第八节 TOXNET 毒理学数据库	101
一、TOXNET 数据库概况	101
二、TOXNET 数据库检索	103
第九节 国际药学文摘数据库	106
一、美国《国际药学文摘》简介	106
二、国际药学文摘数据库的检索	107
三、IPA 数据库与 Medline 的比较	107
第十节 Elsevier 电子期刊全文数据库	108
一、ScienceDirect 数据库简介	109
二、检索运算符	110
三、检索方法及检索结果的显示	110
四、检索结果解读与保存	115

第五章 中文生物医学文献信息资源与检索系统

第一节 中国生物医学文献数据库	116
一、数据库简介	116
二、数据库使用方法	117
三、检索结果的处理	119
第二节 CNKI 系列数据库	120
一、CNKI 系列数据库介绍	120
二、CNKI 系列数据库检索指南	121
第三节 全国报刊索引	127
一、简介	127
二、检索指南	128

第四节 国家科技图书文献中心	133
一、简介	133
二、NSTL 的资源与服务	134
第五节 中国高等教育文献保障系统	140
一、简介	140
二、使用方法	141
第六节 万方数据资源系统	142
一、简介	142
二、检索指南	142
三、数字化期刊资源的浏览和下载(举例)	144
第七节 中国中医药期刊文献数据库	146
一、简介	146
二、使用方法	147
第八节 中国药学文献数据库	148
一、简介	148
二、使用方法	148

第六章 引文信息资源及检索

第一节 科学引文索引	149
一、科学引文索引概况	149
二、引文索引的作用和意义	150
三、科学引文索引使用方法	150
四、科学引文索引的其它数据库	152
第二节 中国科学引文数据库	153
一、概况	153
二、中国科学引文数据库期刊来源	154
三、中国科学引文数据库使用	155

第三节 中国生物医学期刊引文数据库	157
一、概况	157
二、使用方法	158

第七章 美国国立医学图书馆生物信息中心 主要信息资源与利用

第一节 Entrez 系统检索方法与技巧	160
一、Entrez 检索系统的简介	160
二、Entrez 系统的检索	161
第二节 PubMed Central	162
第三节 Genbank (Nucleotide-核苷酸序列数据库)	163
一、Genbank 简介	163
二、Genbank 检索	163
第四节 Genome (基因组数据库)	170
一、Genome 基因组数据库的数据收录范围	171
二、Genome 基因组数据库的主要功能	171
三、Genome 基因组数据库的检索特点	171
四、Genoma 基因组数据库检索结果	172
五、相关数据链接	173
第五节 dbSNP 数据库	173
一、dbSNP 数据库简介	173
二、SNP 数据库的建立和现状	174
三、dbSNP 检索	174
四、SNP 数据库的应用和存在的问题	176
五、中国建立特有 SNP 数据库的必要性	177
第六节 Protein	177
第七节 在线人类孟德尔遗传(OMIM)	180
一、OMIM 数据库的功能	180

二、OMIM 数据库的特点	180
三、OMIM 数据库的查询方法	182
四、OMIM 数据库的缺陷	184

第八章 重要生物医学网站选介

第一节 综合性生物医学网站	185
一、国内综合性生物医学网站	185
二、国外综合性生物医学网站	193
三、综合性医学学(协)会	196
四、综合性生物医学期刊	200
第二节 预防医学网站	202
一、美国疾病控制与预防中心	202
二、美国预防医学会	206
三、泛美卫生组织	207
四、美国卫生和人类服务部	208
第三节 基础医学网站	209
一、欧洲分子生物学实验室	209
二、SWISS-PORT/TrEMBL	209
三、神经科学在线	211
四、微生物学网络	213
五、美国生理学会	214
第四节 临床医学网站	216
一、美国传染病学会	216
二、Jeslin Diabetes Center	217
三、美国心脏协会	219
四、国际 AIDS 保健医师协会	221
五、美国外科医生协会	222
六、癌症预防在研项目目录	224

第九章 网络生物学常用软件工具选介

第一节 PCR 引物设计软件	225
一、PCR 引物设计简介	225
二、引物设计软件 Primer Premier 5.0	228
三、在线引物设计软件 3.0	249
四、primer blast	253
第二节 在线序列比对软件—NCBI-BLAST	254
一、NCBI Blast 简介	254
二、NCBI Blast 的应用	256
第三节 DNA 序列分析软件	258
一、DNA 序列分析软件——DNASTAR 安装	259
二、DNA 序列分析软件 GeneQuest	259
三、DNA 序列引物设计软件 PrimerSelect	265
四、DNA 序列酶切图软件 MapDraw	273

第一章 网络生物医学信息资源概论

第一节 网络生物医学信息资源的类别、特点和评价

信息是事物运动的存在或表达形式,是一切物质的普遍属性,实际上包括了一切物质运动的表征。在人类社会,信息经过组织和排序,能够成为区别于物质和能量的另一种基础性社会资源——信息资源。信息资源广泛存在于经济、社会各个领域和部门,是各种事物形态、内在规律和其他事物联系等各种条件、关系的反映。随着社会的不断发展,信息资源对国家和民族的发展,对人们工作、生活至关重要,成为国民经济和社会发展的重大战略资源。它的开发和利用是整个信息化体系的核心内容。

信息资源有狭义和广义之分:广义的信息资源,指的是信息活动中各种要素(包括信息、信息技术以及相应的设备、资金和人等)的总称。本书里提到的“信息资源”为狭义概念,主要指信息本身或信息内容,即经过加工处理,对决策有用的数据。开发利用信息资源的目的是为了充分发挥信息的效用,实现信息的价值。

传统信息资源的形式为印刷型,主要以纸张或书本作为载体。随着信息存储介质的发展,信息资源可以记录在磁带、磁盘或光盘上。计算机不仅可以存储、处理文本信息,还可以存储、处理声音、图像、动画等多媒体信息。在网络环境下,这些信息资源可以实现存取自由和共享,加快传播的速度。网络信息资源不仅包括各类局域网中信息资源的总和,也指广域网中自由传递与存取的各种信息资源。特别是随着互联网的普及,许多信息资源进入互联网,成为世界各地的用户都可以访问的宝贵网络财富。

网络生物医学信息资源(medical information resources on network)是指以电子数据的形式存储在光、磁等非印刷型的载体中,并通过网络、计算机或终端等方式再现出来的与医学及生物相关的文字、图像、声音、动画等多媒体信息资源。

一、网络生物医学信息资源的类别

(一)按信息内容表现形式和用途分类

1.网络数据库

网络数据库是网络生物医学信息资源的主要形式之一,主要是指出版商和数据库生产商在网络上发行的数据库。它可以是电子报纸、电子期刊、电子图书等一次文献数据库,也可以是索引、目录、文摘等二次文献数据库。

网络数据库经过订购后直接通过 Internet 或以本地镜像站点访问检索,同时依托网络发行传递的方便快捷,将信息检索、原文传递和最新文献报道等服务融为一体。网络数据库主

要包括:题录摘要数据库和全文数据库。

题录型数据库中最为著名的是美国国立医学图书馆(National Library of Medicine NLM)免费提供的 PubMed 数据库, CANLIT、TOXIVET、美国专利数据库、中国期刊题录数据库、中国专利数据库等也都提供免费检索。商用数据库如著名的联机检索系统 Dialog 和 STN 都通过网络提供多种医学数据库查询。OCLC First Search 检索系统提供 3000 多万条文献数据库和书目数据库服务。ISI 的 Web of Knowledge 资源系统提供 Web of Science, Science Citation Index Expanded, BIOSIS Preview, ISI Chemistry 等数据库检索服务。国内有中国生物医学文献数据库(CBMdisc)、中文生物医学期刊数据库(CMCC)及中国科技文献数据库(CSTDB)等文献型数据库,可以对光盘塔局域网或网上注册的用户提供服务。

全文型医学数据库目前主要有 ProQuest Medical Library、ScienceDirect、OVID@journals 全文检索系统、EBSCO 数据库、SwetsNet 全文期刊、Springerlinke 全文期刊、Nature 系列期刊数据库以及中国期刊网(CNKI)提供的清华全文期刊数据库等,这些数据库大都需付费或注册使用。

网络中的数值事实型数据库主要包括基因库、核酸序列、蛋白质结构库等分子生物学数据库,以及毒理学、药物方面的事实型数据库。如美国 NCBI(national Center for Biotechnological Information)提供的 GeneBank、Nucleotide Database、Protein Structure、TOXNET 提供的 HSDB、IRIS 等。网上事实型数据库多为免费检索。

多媒体数据库包括化学物质或药物三维立体结构数据库、生物学图谱库、医学影像库(X 线片、CT 片、磁共振图像)、病理切片库等,如美国国立卫生研究院的“可视人”计划(visible human project)数据库、TOXNET 中的 NCI-3D 和 HSDB 结构库等。

2. 电子出版物资源

电子出版物主要包括电子报纸、电子期刊、电子图书、电子法规等,在网上浏览、订购该类出版物已成为一种发展趋势。与传统印刷型出版物相比具有信息量大、使用更方便,数据更新颖、可靠性更高、承载信息更丰富,具有较强的交互性等特点。

德国 Merck 公司在网上提供默克(Merck)医学信息手册、诊断治疗手册以及药物手册的部分内容,可以免费利用 Free-ebook 网提供包括医学健康在内的免费图书。超星公司对国内 25 万种图书进行数字化加工,并开展网上阅览服务,其中生物医学图书就有上千种。

除了以上介绍的全文数据库外,目前网上免费生物医学期刊主要有斯坦福大学的 Highwire 网站,收录 400 多种医学期刊和电子出版物。此外,一些著名学会的出版物也可通过登录学会网站获取,如美国医学会出版的 JAMA 及系列进展性刊物、美国微生物学会出版的 10 种学会出版物。生物医学期刊出版商也利用网络开展电子期刊的发行,如 BMJ、Proc-Nat-Sci-NSA 等著名的期刊提供免费全文。目前我国部分期刊有自己的网站,提供目次、摘要及部分全文,如生理学报(<http://www.actaps.com.cn>); Cell Research(<http://www.cell-research.com/>); 中国细胞生物学学报(<http://www.cjcb.org/>)等。例如“万方数据资源系统”和“中国学术期刊网”提供数字化的期刊,有近 2000 种科技期刊全文在网上供注册用户使用。

目前上网的报纸已有几千种。各大报纸有自己的网络版,与印刷版报纸相比,有不受时