上海市生物多样性信息管理系统的建立和应用*

赵 斌 唐礼俊 吴千红 陈家宽

(复旦大学生命科学学院生物多样性科学研究所 , 上海 200433)

摘 要 上海市生物多样性信息管理系统是一个在因特网上发布的 Web 数据库 , 用户可以通过浏览器访问 <http://www.ibsfu.fudan.edu.cn> 来进行检索和查询。目前提供的主要数据库包括 2 个有机体多样性本底数据库 : 鸟类资源库、两栖爬行动物及哺乳动物库 ; 5 个专类生物多样性数据库 : 绿化观赏植物库、抗污染植物库、环境污染敏感植物库、害虫害兽库、湿地生物多样性库。同时 , 数据库上网与实现的相关技术也在文中作了详细阐述。最后 , 就目前关于生物多样性信息系统建设中存在的问题发表了一些看法。

关键词 Web 数据库 , 生物多样性 , 信息管理系统 , 上海

The establishment and application of Shanghai biodiversity information management system/ZHAO Bin , TANG Li_Jun , WU Qian_Hong , CHEN Jia_Kuan

Abstract Shanghai biodiversity information management system (<http://www.ibsfu.fudan.edu.cn>) , a web database published on Internet , can be retrieved via a browser such as IE or Netscape. The databases available include two organism biodiversity background database : database of birds , database of Amphibia , Reptilia and Mammalia. Five specific biodiversity databases include : database of anti_pollution plants and pollution_indicator plants , database of ornamental plants , database of pests and harmful animals , database of wetland biodiversity. The design and realization of the system was introduced in the present paper. At last , the author put forward some viewpoints about the problems in the construction of biodiversity information system.

Key words Web database , biodiversity , information management system , Shanghai

Author 's address Institute of Biodiversity Science , School of Life Science , Fudan University , Shanghai 200433

1 引言

保护生物多样性需要各种各样的信息 , 包括人类利用、基础分类、分布、现状和发展趋势以及生态学关系等情况(中国科学院生物多样性委员会 , 1992)。《生物多样性公约》第 7 条特别要求“ 以某种方式组织和维护从查明和监测活动所获得的生物多样性数据”(中华人民共和国濒危物种科学委员会 , 中国科学院生物多样性委员会 , 1997)。近年来信息采集技术和管理技术的新发展使得这些信息变得比以往任何时候更为普及 , 特别是因特网(Internet)技术和 3S 技术的广泛应用已经完全打破了地球上的时间尺度和空间距离。

因特网是目前全球最大的计算机网络 , 它几乎遍及了全球所有的国家与地区。而 Web 则是因特网上最受欢迎、最为流行的信息检索服务 , 它把各种类型的信息资源 , 如文本、静态图像、视频、声音和影像有机的结合起来 , 特别是 Web 数据库把 Web 技术与数据库相结合 , 使用户可以在 WWW 图形用户界面上方便地检索数据库中的内容 , 数据库的存储和管理也上升到了一个新的高度。

上海地区高强度的经济开发 , 改变了许多珍稀物种的的栖息地。作为上海市城市化过程中生物多样性保育研究的基本研究部分 ; “上海市生物多样性数据库建立及保护对策研究”项目组在上海市环境保护局资助

国家自然科学基金重大项目“ 中国关键地区生物多样性保育的研究”(批准号 39893360)和上海市环境保护科技基金(项目编号 : 沪环科 96_011)资助
收稿日期 : 1999 - 12 - 07 ; 修改稿收到日期 2000 - 03 - 10
赵斌 e_mail : zhaobin@fudan.edu.cn

下,经过 2 年多的大量的实地调查和广泛资料分析,取得了丰富的极其重要的基本背景材料和生物多样性变化数据,并在此基础上建立了上海市生物多样性信息管理系统,它现在已经发展成为一个 Internet 上的一个专用网络服务器(<http://www.ibsfu.fudan.edu.cn>),可以提供 24 小时的网络检索服务。它的建立为进一步进行上海市生物多样性保育的研究和生物多样性资源的管理奠定了坚实的基础。

2 系统的设计与实现

本系统采用最先进和最流行的数据库及网络软件,通过自行开发的接口程序,将后台数据库通过 web 方式发布到 Internet 上,力求为用户提供一种格式统一、直观高效、易学易用、功能完备的人机交互图形界面,使得用户在任何地方只要通过能连接 Internet 的计算机上的 WWW 浏览器就可以登录上网进行检索和查询。

2.1 系统的软硬件结构和网络环境

硬件及操作系统 系统的中心服务设备主要由三台高性能 Pentium II 微机组成,分别安装 WINNT Server 4.0 网络操作系统、Unix(Redhat Linux 6.0)网络操作系统以及数据热备份服务器。

网络环境 本信息系统的中心设在复旦大学生物多样性科学研究所网络信息与地理信息系统实验室,网络环境主要是通过中国教育与科技网(CERNET)及国际互联网向国内外用户提供服务。上海市生物多样性信息管理系统及网络结构如图 1 所示。

图 1 上海市生物多样性信息管理系统及网络结构示意图
Fig. 1 Sketch map of Shanghai biodiversity information management system and network topology

2.2 服务器的建立和软件的选择

建立 Internet 服务器 :WINNT Server 内置了建立 Web 服务器的相关工具软件 IIS(Internet Information Server),目前的最高版本为 4.0 ,它可以轻易地建立 WWW 服务、FTP 服务等 Internet 服务,因此作为本系统的首选。同时,在 Unix 机器上还安装了 Apache WWW 服务器,用于提供 Unix 环境下的数据库发布任务。

后台数据库的选择 :对于一个网络数据库 ,后台数据库的选择同样重要。本分布式数据库环境根据具体情况选择了 Microsoft Access 97 和 Sybase for Linux 11.0。Access 数据库界面友好、操作简单 ,便于完成数据的录入、修改(包括更新)和后台管理工作 ,而 Sybase 优良的基于用户的访问也为数据库的安全性提供了可靠的保障 ,同时对于一些敏感数据库还能提供基于不同身份和权限的访问。通过 ODBC(Open Database Connectivity ,即开放的数据库连接)技术 ,我们还能将这些分布在不同环境下的数据库有机地结合在一起。

2.3 软件开发

一般的 HTML 文本展示给客户端(即 WWW 浏览器)的是一种静态的、非交互式的 WWW 界面 ,如果要在 WWW 上提供数据库服务 ,则需要一个接口程序被实时地执行 ,动态地与客户端交换信息。这个程序把客户端的查询请求送到数据库中 ,并将查询结果以 WWW 方式返回给客户 ,所以这里的软件开发主要是用于响应 WWW 页面请求并连接后台数据库的接口程序的开发。考虑到同以前系统的兼容性和以后开发的可拓展性(赵斌 ,蔡庆华 ,1998) ,这个开发工作分别采用了 3 种不同的软件进行 :用 Microsoft Visual Basic 开发 WIN_CGI ,用 ASP(Active server pages)开发的 ASP 页面以及用 PHP 开发的 PHP 页面。

2.4 数据库结构特点

上海市生物多样性信息系统所包含的数据库涉及到许多生物类群 ,而对不同类群的生物进行描述所需要的数据又存在很大差异 ,所以整理这些数据而生成的数据库在结构上具有各自不同的特点 ,这给我们数据采集的标准化带来一定的困难。也就是说 ,这些数据原则上是难于统一的。因此从实际出发 ,在信息系统的建设之初 ,即在进行系统设计、数据收集和整编的时候就没有刻意追求各类群的数据在形式上的统一 ,而是因地制宜地对不同类群的数据库选择最为合适的字段按照其特点进行分析和采集。

一般来说 ,生物多样性的研究包括有机体多样性、遗传多样性、生态系统多样性及人文多样性四个组份 ,相应地生物多样性的研究应该从这四个方面来进行。据此上海市生物多样性信息系统的总体构架也分为这四个部分 ,即有机体多样性子库、遗传多样性子库、生态系统多样性子库及人文多样性子库。本着“收集一部分 ,建设一部分”、“有迫切需要的库最先建立”、“边建设 ,边服务”的原则 ,我们优先建立了有机体多样性子库 ,并从上海市的实际要求出发 ,将这个子库分成本底库(面向从事生物多样性科学研究的专业人员)和专类库(面向大众和其他专业人员)(图 2)。

图 2 上海市生物多样性信息系统(数据库部分)结构图
Fig.2 Structure of Shanghai Biodiversity information system (division of database)

3 主要数据库介绍

以下列出已经建成并上网的数据库,其中3.1~3.2为有机体多样性的本底库,3.3~3.6为专类库。

3.1 鸟类资源库(<http://www.ibsfu.fudan.edu.cn/Database/shdb/nlzy.htm>)

上海市鸟类区系组成随环境变化较大。虽然本地区鸟类研究可追溯到70年前,并留下不少历史记录,但一些种类现在已经难以确定是否存在。因此为了保证本库数据的可靠性,我们选择了3个代表性的研究区域(佘山、浦东和杨浦),并设置大量样地和样带,组织复旦大学教师和聘请外校人员进行了为期2年多的连续观察。与此同时还借助了正在进行的上海市野生动物普查的一些结果,通过对一些历史资料反复核实,最后确定了上海市佘山山林地区、浦东海边滩头及上海市郊区、城区共有鸟类27科142种。其数据项包括:中文名、拉丁学名、目名、科名、数量、季节型、分布、生态特性、保护级别。其中数量分为3个等级:“+++”表示数量多,“++”表示数量一般,“+”表示数量稀少,为稀有种。季节型一栏中,“留”为留鸟,“冬”为冬候鸟,“夏”为夏候鸟。生态特性主要描述了鸟类栖息环境及食性特征。

3.2 两栖爬行动物及哺乳动物(<http://www.ibsfu.fudan.edu.cn/Database/shdb/lqpx.htm>)

与鸟类资源相似,虽然两栖爬行及哺乳类动物有一些历史记录,但人为因素对其影响太大。本库数据的收集主要参照了上海市自然博物馆1991~1993年间所作的全市范围内调查数据,并综合长期野外工作的专家意见和正在进行的上海市野生动物普查部分结果,最后予以确定。它包括上海市两栖爬行动物及哺乳类动物20科49种。数据项为:中文名、拉丁学名、目名、科名、别名、保护等级、形态特征、生态习性及备注等。备注栏内标注有药用、食虫等其它特性。

3.3 绿化观赏植物库(<http://www.ibsfu.fudan.edu.cn/English/Magnolia/Database/ornamental.htm>)

该库总结了上海市园林局编写出版的《上海市园林植物图说》(1980)及编制的“上海城市植物规划名录”(1992),其中包含了在上海有应用前景的植物材料。并根据杨浦区几个公园和校园的植物名录,结合实地调查以及同园林专业人员的认真讨论,最后确定了能代表目前上海观赏植物资源现状的观赏植物119科445种。主要数据项包括:中文名、拉丁学名、科名、生活型、园林用途、生态习性等。植物生活型分为乔木、灌木、藤本、草本。每一项又细分为常绿和落叶两种。园林用途包括绿化树种、绿篱材料、盆栽观赏。观花植物则描绘了花色、花期。生态习性是指温度、湿度和土壤等环境条件要求。同时还在备注栏中标明了植物的药用价值等其它用途。

3.4 抗污染植物和对环境污染敏感的植物库(<http://www.ibsfu.fudan.edu.cn/English/Magnolia/Database/antipollution.htm> 和 <http://www.ibsfu.fudan.edu.cn/English/Magnolia/Database/Indicator.htm>)

本库结合1992年上海市园林局完成的“环保型人工植物群落的研究和实施”项目中总结的上海市抗污染及敏感植物规划名录,根据国内通常采用的抗性及敏感性划分标准,参照国内外所作的有关实验报道进行了重新审定,最后确定了能适应上海气候条件并有相关抗性实验报道的各种抗污染植物87科195种,以及对环境污染敏感的植物55科106种。

其中抗污染植物库数据项包括:科名、种名、适宜环境、根系分布、生长速率、繁殖方式、对 SO_2 、 Cl_2 、HF、Hg、 NH_3 、 O_3 和粉尘的抗性以及对其它污染抗性、绿化用途和其它用途。其中根系分布特征包括:深根、浅根、直根、块根、宿根、鳞茎等。生长速度分为快、中、慢3等。主要繁殖方式分为播种、扦插、嫁接、分株等。抗性等级以数字表示分为5个等级:“0”代表无抗性,“1”代表抗性弱,“2”代表抗性中等,“3”代表抗性中一强,“4”代表抗性强。

对环境污染敏感的植物库数据项包括:中文名、拉丁名、(科名)、生活型、生态学特性、对 SO_2 、 NO_2 、HF、 O_3 、 Cl_2 的敏感等级,其等级分为“强”、“中”、“弱”3等。备注栏中还注有一些具体的试验数据。

3.5 害虫害兽库(<http://www.ibsfu.fudan.edu.cn/Database/shdb/hchs.htm>)

1974年江苏农学院植保系昆虫教研组编写了《农业昆虫鉴定》一书,书中附有主要的农林害虫名录。1976年复旦大学昆虫教研组编写了《农业昆虫学实验》,对上海市主要农业害虫作了详细的描述。1984~1986年上海市园林科学研究所组织人力对上海市园林植物病虫及天敌进行了全面普查。综合这些资料,再参考专家意见,我们确定了上海市园林、农业常见害虫121科886种。数据项中除了中文名、拉丁名、科名外,

还描述了害虫危害植物种类及害兽危害的对象。

3.6 湿地生物多样性库(<http://www.ibsfu.fudan.edu.cn/Database/shdb/shsd.htm>)

该库主要来自华东师范大学 1997 年完成的《上海地区湿地保护对策研究》的研究成果 ,并参照了一些近期的有关湿地生物资源调查报告及研究论文综合而成。该库共收集上海市湿地范围内高等植物 90 种、低等植物 110 种、底栖动物 182 种、浮游动物 134 种、两栖爬行 22 种、湿地鸟类 142 种、鱼类 143 种 ,包括中文名、拉丁文名、(科名)、形态特征、生态习性等。

4 讨论与建议

4.1 当前数据库的新特点

随着计算机技术和信息技术的发展 ,数据库技术也日新月异。在这样的氛围下 ,生物多样性数据库相应出现了许多新的特点。

1)不再是单机数据库 ,而是通过因特网发布的 Web 数据库。Web 数据库具有常规单机数据库无法比拟的优点 :客户端无需配置特别的软件 ,只要通过 WWW 浏览器就可以完成数据库的查询 ;界面友好 ,容易使用 ,通过鼠标的点击就可以完成大部分查询任务 ;系统具有很好的自适应性 ,数据库发生变化时 ,无需更改页面和应用程序 ;Web 可以同时接收多个用户查询浏览 ;不再需要纸张和软盘介质的数据传递和汇总 ,更容易实现真正意义上的分布式数据库。

2)不再是纯文字描述的文本数据库 ,而是具有超链接的多媒体数据库。

3)不仅仅是通常意义上的关系型数据库 ,而是通过因特网实现的分布式数据库。

4.2 有限资源下开发生物多样性信息系统的一个范例

1996 年在全国开展的首次生物多样性信息机构调查表明 :目前既懂生物多样性专业又精通电脑技术的复合型人才十分有限 ,在计算机硬件资源方面 ,微机比较普遍 ,而大型设备略显不足 ;在资金资源上 ,数据管理的资金满足程度仅为 20% ~ 30% ,缺口很大(王长永 ,薛达元 ,1998)。另外 ,在信息系统的计划与投入上 ,往往重视技术支撑系统建设 ,忽视信息系统技术(蔡淇松 ,1999)。本文所设计的信息管理系统能够比较容易在微机上实现 ,略懂计算机的专业人员只需经少量培训就可以从事具体的开发工作。所以本文可以视为在有限的人力资源、资金资源和硬件环境下开发生物多样性信息管理系统的范例。

4.3 数据收集的连续性问题

生物多样性信息管理系统的建设不可能一蹴而就 ,更不可能一劳永逸。国外许多著名的信息系统都是建立在长期和连续的数据观测和收集的基础之上 ,并具有相对统一的规范 ,而国内在这方面相对缺乏。虽然我们能找到许多历史记录和资料 ,但从中我们并不能得到足够规范的数据。同样 ,如果现在从头开始调查、收集和整理 ,则需要多年的数据积累过程和经费支持。因为只有积累了足够长时间的数据的信息系统 ,才能有真正意义上的评价作用 ,否则这种评价至少是片面的或基于猜想的。

致谢 对在建立本信息系统中提供数据和建议的专家 ,参加数据录入的老师和同学表示衷心的感谢 !

参 考 文 献

蔡淇松 ,1999. 关于生物多样性信息系统的问题与建议二、管理与对策. 生物多样性,7(1): 68 ~ 72
王长永 ,薛达元 ,1998. 中国生物多样性数据库管理能力建设. 生物多样性,6(4): 304 ~ 308
赵斌 ,蔡庆华 ,1998. 中国水生生物 Web 数据库的建立. 见《科学数据库与信息技术论文集(四)》. 北京 :科学出版社 ,204 ~ 209
中国科学院生物多样性委员会(编译),1992. 生物多样性译丛(一). 北京 :中国科学技术出版社 ,75 ~ 90
中华人民共和国濒危物种科学委员会 ,中国科学院生物多样性委员会(编译),1997. 生物多样性译丛(四):生物多样性公约指南. 北京 :科学出版社 ,29 ~ 34