

# 普及生物多样性信息学技术 ——澳大利亚植物标本数字图书馆

□ 梁志敏 / 广州少年儿童图书馆 广州 510300

**摘要:** 澳大利亚数字植物标本馆 (Australia's Virtual Herbarium) 是一个免费向公众提供澳大利亚植物标本及相关植物学信息资源的数字图书馆。其资源类型包括数字化标本、图片、出版物等, 涵盖范围广泛, 浏览、检索功能齐全。文章重点从资源组织、技术特征、界面设计、服务特点等方面进行了综合性的评析。

**关键词:** 澳大利亚植物标本馆, 数字图书馆, 植物标本

DOI: 10.3772/j.issn.1673—2286.2013.10.013

## 1 项目概述

“澳大利亚数字植物标本馆” (Australia's Virtual Herbarium, 以下简称AVH) 是由澳大利亚标本馆首脑理事会 (the Council of Heads of Australasian Herbaria, 简称CHAH) 主持的, 属于全球首个实现全国性标本数字化共享的生物多样性数字平台。

CHAH负责为澳大利亚和新西兰的主要标本馆、植物标本馆解决各种行政问题, 增加澳大拉西亚植物标本馆之间的合作和理解, 其业务由宪法和目标声明管辖。其成员包括澳大利亚国家植物标本馆 (阿德莱德)、昆士兰植物标本馆、堪培拉澳大利亚国家植物标本馆、澳大利亚亚热带植物标本馆 (凯恩斯)、塔斯马尼亚植物标本馆 (霍巴特)、墨尔本维多利亚国家植物标本馆、国家植物标本馆 (悉尼新南威尔士州)、北领地植物标本馆 (达尔文)、西澳大利亚州的植物标本室 (珀斯)、艾伦植物标本馆 (基督城, 新西兰)、新西兰惠灵顿



图1 AVH合作机构分布图

蒂帕帕植物标本室、新西兰奥克兰博物馆植物标本馆。

理事会每年举办一次会议, 主办机构在省会城市按以下顺序轮换: 堪培拉, 布里斯班, 墨尔本, 珀斯, 阿德莱德, 新南威尔士州, 达尔文, 霍巴特, 凯恩斯, 新西兰基督城, 新西兰惠灵顿和新西兰奥克兰。

由其管理的AVH是一个具备

搜索功能的包括图片、数字化标本、物种地图等信息资源的数据库, 其资源并非由公众自主上传, 而是由CHAH进行统一收集, 使用统一标准软件对全澳 (现已包括新西兰) 十余家主要标本馆的数百万份标本实施数字化表达, 并通过网络实现标本信息与环境信息的关联表达及共享。AVH实施十多年来, 极

大地推动了澳大利亚植物分类学及生物多样性研究的发展，澳大利亚也由此成为世界上生物多样性信息研究及利用的先驱国家之一。

AVH的网址是http://avh.ala.org.au/，在其网站上有较详细的背景信息，用户也可从数字图书馆首页的搜索引擎和浏览界面中找到资源的数量和来源等相关信息。

## 2 数字资源组织

### 2.1 资源范围与类型

AVH收集、展览的主要对象是植物标本，植物标本是识别植物多样性的关键，它为变化多端的植物及科学研究提供了永久的参考资料。标本还承载和记录着当地的历史，例如一个地区某个时期内的植被种类，或不同时期种类的记载、变化情况，包括一种植物首次引入该地的时间等等。这些标本通常是经过干燥后，小心装订在硬台纸板上，贴上有有关的资料（包括记录采集时间、地点和采集人姓名的采集标签，以及由植物专家鉴定的定名签）。

通过四个多世纪的研究积累，全世界现已认知生物物种约200万种，物种存在的凭证标本达3亿份（件）；我国现已描述和认知的生物物种占世界已知物种总数10%左右，约10万种，凭证标本达3000万份（件）。物种及其存在凭证的重要原始材料是保藏在标本馆中的植物、动物和菌物标本以及植物园中栽培保育的活植物。这些材料是生物分类学、资源学、物种动态的监测、实施保护和和可持续利用，以及其他生物学基础研究最重要的原始材料和基本信息。每个物种和标

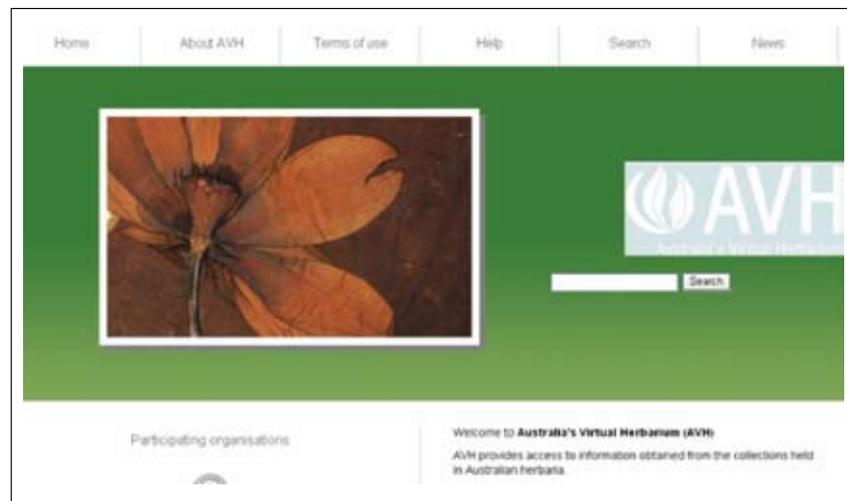


图2 AVH主页

本的描述性信息达上百项，而现代化的实验观测技术，更使每一物种和标本含有的生物信息的信息量呈爆炸性增长。显然，传统的研究方法包括标本室研究方法已无法满足现代信息科学研究及社会大众对认识物种、发现物种和利用物种的迫切要求。

AVH在线展示了包括所有成员馆的标本六百万种，类型有植物、藻类和真菌，从覆盖国家和地区的生物标本基础数据库中获得可信的、新颖的、有效的信息，以能被人们理解和感受的模式表现出来，并使这些原始材料和数据实现资源信息共享。

### 2.2 元数据字段

AVH的每一项馆藏条目都采用统一的元数据方案进行描述，同时使用规范化的受控语言，为揭示条目间的联系提供了线索。

馆藏使用的元数据字段包括题名、目录编号、采集者、采集日期、馆藏地点等（如表1）。在检索页面提供的是资源的题名、目录编

号、采集者、采集日期、馆藏地点、分布地图、收藏分析图表与图片。描述具体细致，增强了标引的准确度，进而提高了检索的检全率和检准率。

## 3 技术特征

### 3.1 界面设计

AVH界面设计简洁清新，功能布局直观便捷，一目了然。

主页主色调为绿色，暗示图书馆的主体资源——植物，分别选用植物油画、标本图片、标本图片资源标识作为题图，并采用幻灯片播放模式进行主动播放，概括性地展现网站主题。

主页设计包括三大模块：主菜单、关键词检索、参与机构介绍。

主菜单占据页面上方，包括六个项目：主页、关于我们、组织团队、帮助、搜索与最新资讯，让用户形成对网站的初步了解，起到导航条的作用。

关键词检索框处于页面正中，在AVH的标志图标之下，非常醒目

表1 元数据字段

Dataset	数据字段	Taxonomy	分类系统	Geospatial	地域
Herbarium	植物标本馆	Scientific name	学名	Country	国家
Catalogue number	目录号	Taxon rank	类目	Locality	地区
Basis of record	记录基础	Kingdom	界别	Minimum elevation in metres	最低海拔
Preparations	保存样式	Phylum	门类	Georeference protocol	地理坐标
Field number	字段号	Phylum	门类	Generalised locality	覆盖范围
Collector	收集者	Order	科目	Verbatim elevation	海拔
Collecting number	收集编号	Family	家族		
Collecting date	收集日期	Genus	种属		
Collection code	收集编号	Species	物种		
Natural occurrence	是否天然	Scientific name authorship	学名作者		
Phenology	物候	Nomenclatural code	命名规范		
Cultivated	栽培	Scientific name without author	无作者的科学名称		
		Infraspecific epithet	俗称		
		Taxonomic issue	分类		
		Name match metric	名称匹配度		
		Name parse type	数据完整度		

地突出网站的主要功能。

“参与机构介绍”栏目处于主页下方，通过超文本连接11个参与机构的主页，通过点击机构标志图标，可进入机构主页，查阅参与机构的相关信息。由于这些机构同时也是AVH数据的提供者，其网站利用鲜艳的图片、生动的生物及景观相片来展示主题，通过浏览这些机构网站，也能激发用户对AVH进行探知的兴趣和积极性。

### 3.2 技术工具及支持

AVH网站是使用JSP语言编写的，固化成HTML页面输出，提高网站的速度以及兼容性。

AVH的数据是由FieldData software组织并存储的，把各地研究机构/人员所做的各项关于澳洲植物研究的成果，收集到由“Atlas

of living Australia”开发并运行的数据库系统中，供网站用户检索。数据是随时更新的，详细记录了标本的地理位置、发现时间等等信息，随着更多的数据汇集，网站内容也越加丰富。

AVH的植物标本附加自制色卡置于遮挡自然光的人工设定室中，采用翻拍灯营造的无影条件下进行翻拍，器材采用数码相机或摄像机。植物标本的形态特征数字化后，作为植物标本图像的原始数据，供计算机进行处理。利用计算机图像处理软件Photoshop等，对原始的植物标本图像进行缩放和旋转、调色、对比度、去噪等后期加工处理，数字图像使用最高分辨率为800dpi，被保存为JPEG格式。最后按照其资料卡对各种植物标本图片资源进行专业标识描述，图片资料卡中详细记录作者、拍照地点、版

权所有、来源标签等描述性数据，其中来源标签一项在数字化后，可点击跳转继续查询。在数字资料卡中还添加了用户互动的问答提问，询问此资源及描述是否正确。

## 4 服务特点

### 4.1 目标用户

AVH面向互联网开放，向所有终端用户提供免费在线资料展览，其目标用户主要分为两类：

(1) 研究植物学、规划植被恢复工作、追踪入侵物种的分布、生物勘探、环境保护规划、资源优化、植物收集和生物多样性调查的用户；

(2) 对植物感兴趣、或是从地球历史角度对植物标本感兴趣的一般大众。

## 4.2 服务方式

### (1) 检索服务

AVH提供五种检索方式:快速检索、高级检索、批目录号检索、类群检索、模糊检索。

快速搜索允许用户快速执行使用一个单一的检索词搜索,对搜索词匹配的第一次尝试是对标准植物名称和通用名称列表。如果没有找到匹配,匹配将尝试对照分类单元名称与标本记录。如果仍然没有匹配,将对以下字段:标本馆名、植物标本馆代码、目录编号、收藏家、收集笔记,进行全文搜索。

高级检索允许用户使用一个或多个检索词进行结构化的查询。用户可以查询的数据字段包括分类名称、分类单元名称等,可在高级检索同时输入最多四个类群的名字进行检索,如果想在超过四个类群名称查询,可使用批次名称搜索,在结果页上添加多个搜索条件。高级检索也提供全文检索项目。

批目录号检索可以通过目录号码进行查询,允许用户执行一个目录号列表的查询。每个目录号码需要在查询框中输入一个单独的行,按此格式,可输入批目录号进行同步查询。批目录号,是不同的植物标本室所用的目录编号格式化存储的数据页的总结。

可以通过植物种属类群名称进行查询。

用户可以进行自定义的模糊检索,允许用户上传ESRI shape文件,并执行一个模糊检索。

检索结果提供三种显示选项供用户查看:记录列表、分布图,以及显示某些方面的统计图表。通过结果页面左侧的菜单中的“选择范围

缩小”选项,并调整结果记录中的显示数值,用户可以压缩原始搜索结果,随着结果范围变小,地图和图表选项卡随之更新。

### (2) 自动完成服务

在高级检索栏中输入时,当用户开始打字,将出现一个分类单元名称建议的下拉列表。这些名字都是ALA的名称列表,提供植物、真菌和藻类唯一的名称,虽然不一定对应AVH中的记录,但ALA名称列表与AVH记录之间的对应关系,将提高全国物种清单的完成度。

### (3) 筛选服务

结果显示页面中“缩小选择范围”的菜单,可以用来申请额外的搜索条件的结果集。每个结果页面的顶部上只列有前三个值,但通过点击“选择更多”选项,用户可以通过一个弹出式窗口中查看所有的结果。如果有超过100个值的结果,只会显示前100个值,但用户可通过滚动窗口浏览接下来的100个值。用户还可以通过点击“标题列值”按字母顺序排列所有未显示的记录数,并通过点击单一的值来过滤其他,也可以使用复选框选择最多15个值,选择这些记录和结果中选定的值,或排除这些记录。

### (4) 个人记录服务

每条标本记录都提供一个总结性的“记录”选项卡。最近期发表(或编辑)的记录将出现在列表顶部(注意,这些不一定是最近采集的标本)。点击个人记录可打开曾浏览记录的详细页面,它提供了所有可用的数据记录。

### (5) 下载服务

AVH所有数据,除某些类群的数据属于敏感保护或生物安全方面的原因,是向所有人开放的。用户可以将搜索结果中的记录,或独特的

分类名称中搜索出的结果罗列成一个物种清单,以电子表格的形式进行下载。

### (6) 下载地图服务

通过点击“下载地图”按钮,用户可以下载一个高分辨率的地图。在弹出的窗口中,用户可以选择格式、大小和分辨率的地图基础层的大小、颜色和不透明度的地图符号。

### (7) 提醒服务

当AVH出现与用户搜索条件相匹配的新纪录时,或用户搜索结果记录中出现由用户贡献的新注解时,用户可以使用代理功能得到电子邮件提醒通知。用户只需要登录订阅电子邮件提醒,便可以选择每月、每周、每天或每小时的基础上接收提醒。用户可以选择搜索结果的任何形式更新条件,如产生新记录、注释或图像添加到AVH时接收提醒。

## 4.3 版权管理

除了AVH的标志,以及另有注明的项目外,AVH网站上的所有记录提供澳大利亚署名的知识共享许可证,表示所提供的数据源是公共的,允许重新分配基于所述数据的任何衍生作品。网站提供链接访问的相关网站,均持有法人代码为CC 3 AU许可证。

同时AVH也表明用户使用AVH所提供的数据时风险自行承担。

## 4.4 用户隐私保护

AVH使用谷歌分析监测用于统计网站的使用,包括用户的互联网协议(IP)地址和顶级域名,网站访

问的日期和时间,访问页面,查询执行和数据下载,以前的网站访问和使用的浏览器等信息。任何未经授权的访问此系统是禁止的,受到澳洲的刑事和民事法律保护。

用户从AVH中访问或下载数据不需注册或登录,但在使用一些额外功能时,需要用户注册登记,如需要登录才能提交问题与AVH记录和注册电子邮件提醒。注册时需提交一个真正的电子邮件地址,在某些情况下,还可能会要求用户提供真实姓名或其他识别或联系的细节,但这样的个人信息将按照ALA隐私政策管理。

## 5 评价与建议

植物标本是识别植物多样性的关键,它为变化多端的植物及科学研究提供了永久的参考资料。标本馆既是自然历史档案馆,也具有图书馆的功能。计算机技术及网络信息技术的发展使得传统标本馆

成为生物多样性的重要信息库。生物标本及其标签信息的电子化、信息化使得“数字植物标本馆”快速发展。AVH对于保护植物、研究生物多样性信息是一个很有意义的项目,其中提供的图片、记录、地图资源可获性及有用性极高,使得植物标本馆从传统的以标本借阅为主的单功能服务向以信息和知识收集、传播发布、检索为主的多功能服务转化,从“被动式”服务转向“主动式”服务。其特点是收藏数字化、传递网络化、资源共享化和结构连接化。在设计方面,界面较为友好,检索功能强大,响应速度快,检全率和检准率高,并能按用户要求提供各种检索及筛选服务,地图下载、提醒等服务具有创新意识,其采用的FieldData software很好支持了该数字图书馆资源的浏览分享系统,为用户提供了更为便捷的获取数据资源的途径。版权管理及用户隐私保护意识强烈,值得学习。

AVH作为功能性的网站,在一

般用户友好度的设计上也存在不足的方面:

(1) 资源推荐力度不够。虽然网站会根据旧用户的浏览记录推荐相关记录,但对新用户是没有任何的推荐与提示的,可以对热门或者更新的资源添加“Hot”或“New”标签,方便新用户使用。

(2) 对一般用户而言界面友好度不足。网站设计的初衷是面向研究植物学、资源优化、植物收集和生物多样性调查等的专业型用户,故而网站进入资料库的主要途径为搜索引擎,对于一般的兴趣用户,则缺乏一个方便快捷的浏览途径。可在主页添加数据库名录的连接,或植物科属的词典连接,方便一般用户点击进入进行普及性的浏览。

(3) 用户互动与分享性不足。可设计论坛等功能模块,供用户相互交流,也可增加更多的社会化分享服务,如微博或电邮分享等,加大对馆藏资源的推广。

### 参考文献

[1] Australia's Virtual Herbarium (AVH) [EB/OL]. [2013-04-15]. <http://avh.ala.org.au/>.

### 作者简介

梁志敏 (1978-), 女, 馆员, 学士, 广州少年儿童图书馆工作, 研究方向: 阅读推广、文献资源建设等。E-mail: hazyliaang@163.com

### Popularity of Biodiversity Informatics Technology – Australia's Virtual Herbarium

Liang Zhimin / Guangzhou Children's library, Guangzhou, 510300

Abstract: Australia's Virtual Herbarium is a free digital library which provides local related botany herbarium digital information resources, including digitized specimen, photographs, publications, etc., covering a wide range with complete functions of browsing and searching. This article comprehensively analyses its resource organization, technical features, interface design, service characteristics and other aspects.

Keywords: Australia's Virtual Herbarium, Digital library, Plant specimens

(收稿日期: 2013-06-18)