

学校编号: 10384

分类号 _____ 密级 _____

学号: 21720081152603

UDC _____

厦 门 大 学

硕 士 学 位 论 文

基于 Flex 与 J2EE 的中国山茶科植物鉴定系统

**The System For Identification Of The Theaceous Flora
Of China Based On Flex Integrating J2EE**

张佳鸿

指导教师姓名: 侯学良 副教授

专 业 名 称: 细胞生物学

论文提交日期: 2011 年 4 月 30 日

论文答辩日期: 2011 年 6 月 02 日

学位授予日期: 2011 年 月 日

答辩委员会主席: 李振基教授

评 阅 人: _____

2011 年 6 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（）课题（组）的研究成果，获得（）课题（组）经费或实验室的资助，在（）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

目录

摘要	1
Abstract	3
第一章 绪论	5
1.1 检索表概述	5
1.1.1 检索表定义	5
1.1.2 传统检索表	5
1.1.3 多通道检索表	6
1.2 数据库在植物分类学的应用	9
1.2.1 生物多样性信息系统	9
1.2.2 植物形态特征数据库	10
1.3 山茶科植物概述	10
1.3.1 山茶科的形态特征	10
1.3.2 山茶科系统学研究概述	11
1.3.3 山茶科植物数量及分布	12
1.3.4 山茶科植物资源及经济价值	13
1.4 研究目的及意义	14
第二章 研究内容及方法	16
2.1 研究内容	16
2.2 研究方法	16
2.2.1 信息收集	16
2.2.2 数据分析及图文处理	17
2.2.3 信息系统构成	17
2.2.4 技术路线	17
第三章 山茶科植物性状研究	19
3.1 分类性状类型的分析	19
3.2 性状的处理	19
3.3 山茶科植物生活型描述	20
3.4 山茶科植物营养体性状	20
3.5 山茶科植物花性状	20
3.6 山茶科植物果实性状	22
3.7 性状描述数据统计	23
第四章 系统构建的相关技术	24
4.1 数据库技术	24
4.1.1 关系数据库	24
4.1.2 SQL 语言	24
4.1.3 MySQL 数据库	25
4.2 J2EE 简介	25
4.2.1 J2EE 概述	25
4.2.2 Spring 框架概述	26
4.2.3 iBATIS 框架概述	26

4.3 Flex 技术	26
4.4 Flex 与 J2EE 的集成	27
4.4.1 Flex 与 J2EE 的集成方式	27
4.4.2 Flex 与 J2EE 的通信方式	28
第五章 系统的设计与实现	30
5.1 系统功能分析	30
5.2 系统数据库的设计	31
5.3 系统架构的实现	33
5.3.1 持久层的实现	33
5.3.2 业务逻辑层的实现	37
5.3.3 表示层的实现	40
5.4 系统功能的实现	42
5.4.1 登录模块	42
5.4.2 信息管理模块	42
5.4.3 鉴定模块	46
第六章 总结与展望	49
6.1 主要结论	49
6.2 展望	50
参考文献	51
附录 中国山茶科植物性状描述	57
致谢	79

CONTENT

ABSTRACT(in Chinese)	1
ABSTRACT(in English).....	3
Chapter 1 Introduction	3
1.1 General of key.....	5
1.1.1 Key.....	5
1.1.2 Traditional key.....	5
1.1.3 Multi-access key.....	6
1.2 Plant database	9
1.2.1 Information system in biodiversity.....	9
1.2.2 Plant morphological database	10
1.3 Overview of theaceous flora of China.....	10
1.3.1 Morphology of theaceous flora of China.....	10
1.3.2 Systematic study of theaceae	11
1.3.3 The amount and distribution of theaceous flora.....	12
1.3.4 The value of the theaceous flora.....	13
1.4 Significance of this study	14
Chapter 2 Area and methods of study.....	16
2.1 Area of study.....	16
2.2 Methods of study.....	16
2.2.1 Collection of information.....	16
2.2.2 Analysis of data	17
2.2.3 Establishment of system	17
2.2.4 Roadmap of technology	17
Chapter 3 Character description of theaceous flora of China	19
3.1 Analysis of character type	19
3.2 Research about character.....	19
3.3 Description of life form.....	20
3.4 Description of vegetative mass	20
3.5 Description of flower	20
3.6 Description of fruit.....	22
3.7 The statistic of character descriptor	23
Chapter 4 Technology.....	24
4.1 Database	24
4.1.1 Relational database	24
4.1.2 SQL.....	24
4.1.3 MySQL	24
4.2 J2EE framework	25
4.2.1 J2EE.....	25
4.2.2 Spring.....	26
4.2.3 iBATIS	26

4.3 Flex.....	26
4.4 Intergration of Flex and J2EE	27
4.4.1 Intergration mode between Flex and J2EE	27
4.4.2 Communication between Flex and J2EE	28
Chapter 5 Design and implementation of the system	30
5.1 Analysis of demand	30
5.2 Design of database	31
5.3 Implementation of system's architecture.....	33
5.3.1 Implementation of persistence tier	33
5.3.2 Implementation of business logic tier.....	37
5.3.3 Implementation of presentation tier.....	40
5.4 Implementation of system's function.....	42
5.4.1 The module of login.....	42
5.4.2 The module of information	42
5.4.3 The module of identification.....	46
Chapter 6 Conclusion and Prospect.....	49
6.1 Conclusion	49
6.2 Prospect	50
Referance	51
Appendix The table of character description for theaceous flora of China	57
Acknowledgement.....	79

摘要

随着分类学的发展,传统检索表已经不能满足人们的需要。传统检索存在的一些固有缺点,使得人们需要花费大量的时间与劳动力用于检索工作,这极大的影响了分类研究的工作效率,因此人们急需一种更为有效的检索工具。计算机强大的数据处理能力和海量的数据存储能力为发展一种更为有效的检索工具提供了技术支持。利用计算机产生的交互多通道检索表与传统二歧检索表相比具有众多的优势。

本研究以中国山茶科植物为例,利用植物志等现有资料,结合标本考证,建立中国山茶科植物形态特征数据库。然后,利用 Flex 与 J2EE 技术构建在线交互多通道鉴定系统。主要研究成果如下:

通过性状研究共总结出中国山茶科植物性状描述符 155 项,其中质量型描述符 106 项,数值型描述符 49 项。性状描述类型共分为 6 大类:生活型(1 项)、营养体性状(25 项)、花的性状(100 项)、果实的性状(26 项)、物候(2 项)、特殊性状(1 项)。总体上来看,抽取出来的营养体性状明显比花与果实的性状少。但是花与果实只有在特定时期才有,可获得的周期短,然而营养体的一些器官在植物的整个生命周期都存在,因此在多通道检索过程中,用营养体性状进行鉴定更具优势。

在建立中国山茶科植物形态数据库的基础上,本系统采用了现在流行的 Flex 技术和 J2EE 技术,以 Flex 构建客户端,利用 Java 技术来处理后台数据,成功建立了基于网络的中国山茶科植物鉴定系统。

整个系统分三个模块:登录模块、信息管理模块和鉴定模块。每个模块都有各自的功能:(1)登录模块负责用户权限的控制。只有拥有系统管理员权限的用户才能使用信息管理,而普通用户只能通过浏览进入鉴定模块。权限的区分保证了整个后台数据库的安全。(2)信息管理模块为系统管理员提供对后台数据库进行修改的平台。通过使用这个模块,用户可以在线完成对物种名录、图片、分类性状信息及物种性状信息等信息的管理。(3)鉴定模块是整个系统的核心,它提供了在线交互鉴定功能。鉴定结果是以物种符合鉴定条件的个数进行排列的,即

使某一物种仅仅符合众多鉴定条件中的其中一个，也会被列出来以供参考。这种做法尽可能地减少鉴定所产生的错误。同时，用户可以很便捷地浏览物种的性状信息及图片。

本系统的特点：（1）较强的鉴定功能。系统建立了完善的形态数据库，客户端采用 Flex 开发，虽然运行于浏览器中，但运行界面接近于桌面的运行程序，同时通过对检索的优化，能够减少鉴定的误差，交互及时性强，准确率高。（2）易维护，易扩充。从数据内容上来看，由于本系统提供了独立的在线编辑工具，可在任何有网络的电脑上进行编辑，而不必要安装任何软件。从系统架构上来看，本系统采用了 Flex+Spring+iBATIS 多层结构设计，各层之间相对独立，因此对每一层进行维护是容易的，维护成本低。（3）可移植性。本系统是基于 Flex 与 J2EE 建立的，由于 Flex 和 J2EE 技术都与平台无关，因此可以在不同的操作系统之间进行切换。

关键词：中国山茶科；多通道检索表；Flex；J2EE

Abstract

With the development of the taxonomy, the traditional key cannot meet the demand any more. People have to spend a large amount of time and labor force in retrieving due to some of the drawbacks lying innately in the traditional key, which has a great influence on the work efficiency of the taxonomy and urges people to find a search tool with higher efficiency. Computer, together with its powerful data-handling capacity and the mass data-storage capacity, provides technical supports for the development of the search tool. The interactive multi-access keys are superior to the dichotomous keys in numerous aspects.

Using existing data such as the flora and integrating with the specimens research, the following research is illustrated by the example of the theaceous flora of China to set up a morphology database. The system of online interactive multi-access identification is builded based on this database. The main research results are as follows:

It is concluded from the character study that there are a sum of 155 character descriptors of the theaceous flora of China, including 106 multistate character descriptors and 49 numeric character descriptors. The character descriptors are divided into 6 types: the life form (1), the character of vegetative mass (25), the character of flower (100), the character of fruit (26), the phenological phenomena (2), and the special character (1). Overall, the extracted number of the character of vegetative mass is much less than the character of flower and fruit. The flower and fruit exist in the specific period with the short cycle of accessibility, while some organs of the vegetative mass exist in the whole life span of the plants. For that reason, it is better to carry out the identification with the character of vegetative mass during the process of multi-access retrieval.

Based on the database for the theaceous flora of China, with the Flex to build the client and the Java technology to process the backend data, the following system adopts the prevailing Flex technology and J2EE technology to successfully set up the identification system based on the Internet for the theaceous flora of China. As a result, it brings the on-line data management and identification for the the theaceous flora of China into being, and make full use of the features of the Internet.

The whole system is divided into three modules: the module of login, the module of information, the module of identification. And each module has their respective functions: (1) the module of login takes responsibilities for the administration of the user permission. Only those users who have system administrator privileges can apply the information management, while the common users can but enter into the module of identification through browsing. The safety of the background database is guaranteed by the differentiation of the privileges. (2) The module of information provides a platform for the system administrator to modify the background database. By applying this module, the users can administrate the information of the species

on-line, such as the basic information, pictures, character information, etc. (3) The module of identification is the core of the whole system and it provides the on-line interactive functions. The appraisal results are ranked according to the numbers of the species fitting the appraisal conditions. Even though a species just fits in with one of the numerous appraisal results, it will be listed for reference. This method cuts down the errors made by the identification as far as possible, meanwhile, it enables the users to browse the character information and pictures of the species conveniently and fast.

The characteristics of this system are presented as following:

(1) The system has a stronger function of identification with the establishment of the thorough database for morphology. The client is developed by employing the Flex and operated in the browser, but the operation interface is similar to the desktop application. Meanwhile, with the optimization of the retrieval, it will reduce the retrieval error, make the interaction more prompt, and have a high precision rate.

(2) The system is much easy to be maintained and expanded. In terms of the data content, due to the provision with the independent, on-line editing tool by the system, it can be operated in any computer accessing to the network without installing any software. In terms of the system architecture, with the application of the multi-layer structure design of Flex+Spring+iBATIS and the relative independence between layers, the maintenance for every layer is easier to carry out with low cost.

(3) The system has the portability. Established on the Flex technology and the J2EE technology, which are independent from the platform, it can be switched in the different operating system.

Keywords:Theaceae; multi-access keys; Flex; J2EE

第一章 绪论

1.1 检索表概述

1.1.1 检索表定义

生物检索表(key)是以鉴定与区分生物为目的而编制的表,它是分类学研究的重要工具,能够帮助使用者快速地鉴定未知植物,是构成植物志、手册、专著及其他形式分类学文献的重要组成部分^[1]。植物分类学不仅是植物学的基础,也是生物学的基础,同时它与农、林、牧、渔、中医药及化工等行业也有着密切的联系^[2]。

1.1.2 传统检索表

检索表最早的雏形可以追溯到希腊时代的草药学专著,编者尝试按照二叉分类法将草药归类,这在形式上与现在的检索表很相似,但是由于它并不是基于科学分类法,因此在本质上与现在的检索表是不同的^[1]。法国生物学家拉马克(Lamarck)在 1778 年首次应用的二歧检索表(dichotomous keys)是真正意义上的检索表^[3]。二歧检索表是发明最早的、目前最经常使用的检索表,也称为传统检索表。

1.1.2.1 传统检索表编制原则及种类

传统检索表就是依据拉马克二歧分类原则编制的,即选用明显而又相关的形态特征把同一类群的生物分成相对应的两个分支,再把每个分支中的分类群再用相对的形态特征分成相对应的两个分支,依次下去,直到将所有分类群分开为止^[4]。

使用传统检索表检索的时候,只有一个起始点,遇到成对性状组成的分支时,只能往某一个分支进行检索,因此整个检索过程是沿着固定的顺序进行的,所以同时也称为单道检索表(single-access keys)。

传统检索表按照所包含分类群的范围可分为分科检索表、分属检索表、分种检索表等;按照编制的依据可分为人为检索表与系统检索表;按照编排的形式分

为定距式检索表、平行式检索表和连续平行式检索表^[1,2,5]。其中定距式检索表是最为常见的一种检索表。

1.1.2.2 传统检索表的不足

生物分类鉴定时使用最多的检索表是传统的二歧检索表^[6]，但是这种检索表存在着诸多的缺点^[7-11]：编制者需要将要编制在检索表内的类群进行全面而详细的研究；编制过程中为了遵循二歧分类原则，不得不放弃一些可用的有价值的性状；使用者进行检索时可能遇到检索表上需要的性状，然而由于标本所处侯物期等原因而无法观察到的性状；使用者对某一个性状观察不准确，导致无法获得检索结果等等。

1.1.3 多通道检索表

由于传统检索表存在着一些固有缺点，人们寻找到一种更有效的检索表——多通道检索表(multi-access keys)。多通道检索表与单道检索表是相对应的，使用单道检索表检索某一物种时，只能沿着固定的顺序进行，在每一个分支只能选择一对相对性状的其中一个；而多通道检索表以使用者为中心，允许使用者自己决定性状的使用顺序，在每个分支可以从多个性状中选择自己想用的性状(如图 1-1)^[12]。

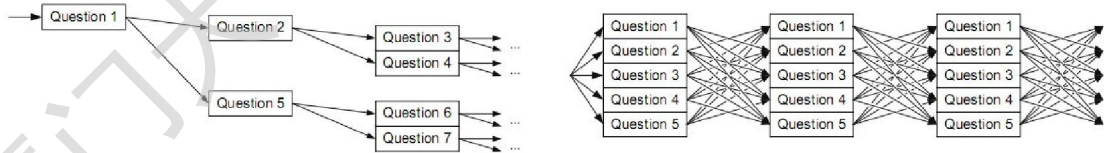


图 1-1 单道检索表(左边)和多通道检索表(右边)使用顺序^[13]

Fig. 1-1 User interaction steps in a single-access key (left) and a multi-access key (right)

1.1.3.1 多通道检索表早期的形式

最早的多通道检索表是以穿孔卡片的形式出现的^[14]。可是，这种形式的多通道检索表并没有充分利用数据，当使用这种检索表检索出一个符合一系列性状组合的分类群时，用户获得的关于这个分类群的性状信息仅限于卡片上所列举的

[11]; 再者, 由于卡片的限制, 所能列举的性状也是有限的, 对于物种的增减也是极为不方便的。这种检索表的实现方式较为繁琐, 因此没有得到大规模的应用。

1.1.3.2 计算机辅助下的检索表

随着计算机的迅速发展, 特别是编程技术、多媒体技术和网络技术的普及, 生物分类鉴定方法得到了长足的进步^[5], 科学家们研究出了许多新的分类鉴定工具, 主要包括两类, 一类用于产生最佳的二歧检索表, 并打印出来使用; 一类是交互式鉴定工具。前者几乎没有多少进展, 而后者得到了很大的发展, 变得多种多样^[12, 15]。

交互式的鉴定工具主要分为交互式检索表(interative keys)、专家系统(expert system)、神经网络(neural networks)和计算机视觉识别(computer vision in species identification)^[12, 15, 16]。其中发明最早的, 应用范围最广的是交互式检索表。

交互式检索表包括两种类型: 交互式二歧检索表(interactive dichotomous keys)和交互式多通道检索表(interactive multi-access keys), 其中交互式多通道检索表得到更为普遍的使用^[17]。因为交互式二歧检索表只是传统纸质二歧检索表的电子化, 并没有改变二歧检索表所固有的缺点, 而交互式多通道检索表克服了早期多通道检索表的一些缺点, 让使用者能够更加高效地应用这种工具进行鉴定。

1.1.3.3 交互多通道检索表工具

DELTA 系统可能是出现的最早, 现在使用范围最广的电子检索工具^[18-20]。它是由澳大利亚联邦科学工业研究院(CSIRO)昆虫研究所 Dallwitz 等人设计、建立并逐步完善的信息处理系统^[20]。DELTA 系统包含了一系列的应用程序, 能够执行诸多功能: 分类语言的自动生成、传统二歧检索表的生成、交互式多通道的检索、表征分类运算所需数据格式的转换等等, 其中最吸引人的功能就是交互式多通道检索 DELTA-Intkey^[19]。

后来的研究人员, 为了使 DELTA 系统更好地应用于各类平台, 在 DELTA 系统的基础上, 发展了各种单机的鉴定工具, 包括 DeltaAccess、DIANA、PANKEY 等^[21]。

由于 DELTA 系统对于网络的支持程度不够, 后来又陆续出现在了以 DELTA

系统为基础的各类应用。NaviKey^[22]是使用 DELTA 数据格式的 Java applet 的应用,其能够直接嵌在网页中使用,但是对于数据的编辑仍然需要 DELTA 的支持。Brach 等^[23]研发的 ActKey 是一个基于网络的多通道检索程序, 后台服务器采用 Java 语言编写, 是一个典型的动态网页应用。

另一个广泛使用的就是 Lucid 家族产品, 包括 Lucid Professional 和 Lucid Phoenix, 它们是由澳大利亚昆士兰大学有害生物信息技术与推广中心(CPITT)研究开发成功的^[24, 25]。这两个软件都是商用的, 其中 Lucid Phoenix 主要是用来编写和使用基于计算机的二歧检索表^[26], 而 Lucid Professional 是基于多媒体技术的多通道分类检索系统, 它所产生的检索表可以嵌在网页中使用, 目前最新的版本是 Lucid V3.5, 使用是要付费的, 而 Lucid V3.3 现在可以免费使用^[27]。

Xper2 是用 Java 平台开发的多通道检索软件, 有自己的格式 Xper, 同时也可以采用 SDD 格式, 目前不支持数值类型的检索^[28]。3I 是基于 ASP 动态网页与 MS ACCESS 数据库的网络多通道检索系统, 它对于浏览器的使用有所限制^[29]。其它工具还有 X:ID、SLIKS、ETI LinnaeusII/III 等。

目前, 交互多通道检索表的研究与开发主要集中于国外^[30]。我国在这方面的研究比较滞后, 最早的研究见 20 世纪 90 年代^[31, 32], 但由于结构比较, 内容也欠丰富, 并没有形成成熟应用的产品。近年来新开发比较成熟的应用主要有, 高灵旺等^[33, 34]的 TaxoKeys、张小斌^[35]的 InsectID 等, TaoKeys 是单机版的, 而 InsectID 是针对于昆虫的, 而且需要付费。

1.1.3.4 交互多通道检索表的优点

交互多通道检索表具有诸多的优点^[30, 36, 37]:

(1) 不限制检索性状的使用, 使用者可以忽略难以判断或者标本不存在的性状, 而传统的二歧检索表由于检索路线是固定的, 在检索的过程中, 只能按照检索表所提供的性状进行检索, 一旦遇到标本不存在的检索性状时, 往往很难继续进行检索;

(2) 在检索的过程中, 如果发现问题, 使用者可以随时将已经使用的部分性状移除, 或者改变一个已经选择的性状属性, 这一点体现了计算动态处理数据的优势, 节省了检索所花费的时间;

(3) 在检索的过程中, 可以直接使用数值型性状, 而不是将数值划分为小范围, 然后转存为质量型性状形式;

(4) 容易对数据进行更新和维护等等。

1.1.3.5 目前存在的不足

目前已经有很多工具能够支持网络检索, 但是对于数据的编辑仍然要依靠其它单机软件, 无法发挥网络随时编辑的优势。同时, 由于数据是依靠本地计算机进行编辑的, 造成了网络检索时对图片的利用不足, 无法在线获得足够的图片。

1.2 数据库在植物分类学的应用

1.2.1 生物多样性信息系统

生物物种资源和植被资源是人类社会赖以生存的重要物质基础, 是自然资源和环境的重要组成部分^[38]。随着计算机技术及国际社会对生物多样性信息收集和管理重视程度的提高, 目前世界各国及国际组织已经建立了多种多样的生物多样性信息系统^[39]。综合性的多样性信息系统, 主要有物种 2000 年(Species 2000)、世界植物名称索引数据库(IPNI)、生物多样性保护信息系统(BCIS)、中国科学院生物多样性信息系统(CBIS)等; 区域性和专属的多样性信息系统, 主要有

“Invaders Database System(美国东北部外来植物数据库系统)”、“Gymnosperm Data-base(德国裸子植物数据库)”、“江苏植物资源信息系统”、“云南省珍稀濒危植物信息系统”、“梁子湖湿地生物多样性信息系统”等^[38-44]。

生物多样性信息系统是以数据库技术为核心, 将生物多样性信息和数据储存在电子计算机的贮存器上, 在计算机硬件与软件的支持下实现对生物多样信息和数据的保存、更新、保密、检索、统计、评价、预测、模拟和决策等功能的应用系统^[45, 46]。

在整个生物多样性信息系统中最基础的内容就是数据库。现在随着计算机各种技术的成熟, 可以利用数据库作为知识源, 把逻辑学、模糊学和统计学等多个学科的成果综合应用, 从而实现对库中数据的充分利用, 获得一些与数据库数据相吻合的宏观信息^[43, 47]。因此, 提高数据库的完善程度, 再结合系统中的其它工具, 能够在一定程度上提升整个系统的综合应用。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库