

《天文学名词》数据库新版网站设计与实现^{*}

崔辰州¹, 余 恒², 卞毓麟³

(1. 中国科学院国家天文台, 北京 100012; 2. 北京师范大学天文系, 北京 100875; 3. 上海科技教育出版社, 上海 200235)

摘要:《天文学名词》数据库网站是天文学名词审定委员会开展名词审定工作和为广大用户提供名词查询检索以及其它相关服务的网络平台。随着名词审定工作的不断推进和广大用户对名词数据库功能需求的不断增加, 原有的网络平台从功能和信息量上都无法满足当前的需求。《天文学名词》数据库新版网站借助成熟的数据库和网络技术, 围绕名词的征集、审定、发布、查询检索等环节, 充分实现了广大用户和天文学名词审定委员会所需要的功能, 提供了更加丰富的信息。介绍了新版数据库及网站的设计原则、实现方法, 并从功能需求出发着重介绍后台数据库和用户界面的设计与实现。《天文学名词》数据库网站是一个典型的由数据库驱动的简单网络应用系统, 本文所涉及的工作和介绍的经验将为类似应用系统的开发提供切实可行的参考。

关键词: 天文学名词; 需求分析; 数据库设计; 网站设计; 程序开发

中图分类号: TP393.092 **文献标识码:** A **文章编号:** 1672-7673(2011)02-0178-07

1983年, 天文学名词审定委员会成立^[1], 当时是中国天文学学会的一个工作委员会。1985年全国自然科学名词审定委员会(现全国科学技术名词审定委员会, 以下简称“全国名词委”)^[2]成立, 下设天文学名词审定委员会(以下简称“天文学名词委”)。从此, 天文学名词工作正式纳入国家名词审定体系, 天文学名词委接受全国名词委和中国天文学学会的双重领导。

天文学名词委成立以来完成了大量的工作, 取得了一批学术价值和社会效益都非常良好的研究成果。

1986年, 天文学名词委完成首批1956条基本名词的审定, 并上报全国名词委。1987年7月10日, 国家科委和中国科学院向国务院呈送《关于公布天文学名词的请示》。1987年8月12日, 《国务院关于公布天文学名词问题的批复》(国函【1987】142号)中指出: “全国自然科学名词审定委员会是经国务院批准成立的。审定、公布各学科名词, 是该委员会的职责范围, 经其审定的自然科学名词具有权威性和约束力, 全国各科研、教学、生产、经营、新闻出版等单位应遵照使用。请告全国自然科学名词审定委员会, 天文学以及以后各学科经审定的自然科学名词, 国务院授权该委员会自行公布。”1987年9月, 由天文学名词委审定、全国名词委公布的《天文学名词(1987)》一书由科学出版社出版。在所有学科中, 这是首先完成的。此后, 又相继完成出版了《天文学名词(1987)》(海外版)、《英汉天文学词汇》(第二版)、《汉英天文学词汇》、《英汉天文学名词》、《天文学名词(1998)》等辞书。当前, 《海峡两岸天文学名词》的编订工作已基本完成, 即将出版。

上世纪90年代中后期互联网开始进入普通中国人的视野, 并快速发展成为人们日常生活、学习和娱乐的主流平台。曾任天文学名词委主任、北京天文台台长的李启斌先生, 早在2000年就意识到互联网对名词工作的重要性, 提出要尽快利用互联网来进行科技名词的推广应用工作^[3]。

从2002年开始, 天文学名词委开始利用网络和数据库等技术手段为广大用户提供名词服务^[4]。

^{*}基金项目: 国家自然科学基金(60603057; 10778623; 10820002; 60920010; 90912005), 国家高技术研究发展计划(863计划, 2006AA01A120), 北京市科技新星计划(2007A085), 中国科学院“十一五”信息化专项(INFO-115-C01-SDB3-04)。

收稿日期: 2010-04-26; 修定日期: 2010-05-26

作者简介: 崔辰州, 男, 副研究员, 研究方向: 虚拟天文台、天文信息技术, Email: ccz@bao.ac.cn

最初阶段,通过把2000年由上海科技教育出版社出版的《英汉天文学名词》电子版正文部分导入到一个小型数据库,开发了“英汉双向天文学词典”,提供基本的词条检索功能^[5]。

多年来,《天文学名词》的数据库和网站一直维持着一种简单、基本的形式。当今世界,包括天文学和信息技术在内的各类科技都在日新月异地发展着,科技新词快速涌现,学术交流日益频繁,知识全球化不断深入,信息沟通和共享越来越方便快捷。这些变化和发展都对多年来基本维持不变的《天文学名词》数据库网站提出了挑战:

(1) 天文学名词委需要快捷、高效的网络平台和工作环境进而提高审定效率以应对越来越多的新词;

(2) 两岸三地学术交流的日益频繁不但需要英汉对照,还需要中文繁体和两岸译名对照等更多信息;

(3) Web 2.0、应用服务化(XaaS)等互联网新兴发展潮流需要名词数据库提供更丰富的接口和功能,以满足各种类型用户和客户程序的需要。

为了满足上述需求,丰富数据库的内容和服务形式,及时利用新兴的技术成果,天文学名词委决定对原有数据库和网站系统进行全新改版。

《天文学名词》数据库及网站是一个典型的由数据库驱动的简单网络应用系统。虽然在新版的设计与实现中并不存在什么技术上的挑战和全面创新,但鉴于目前国内天文学研究和科普教育领域对数据库和网络技术应用水平还很落后,这样一个已经成功运行多年并广受科研人员和用户好评的系统必将为类似应用系统的设计开发提供切实可行的参考,促进信息技术在天文学领域中的应用,提升应用的水平。

本文从对新版数据库网站定位和设计原则的介绍入手,然后进行需求分析,第3部分阐述依据需求分析设计的用户界面,第4部分介绍后台数据库的设计,第5部分介绍具体开发过程中的一些技巧并对新版上线运行的效果进行评价。最后对尚未实现的一些功能以及未来发展展开讨论。

1 定位和设计原则

《天文学名词》数据库新版网站的性质是天文学名词委的官方网站,网站的定位是天文学名词委为广大用户提供天文学名词服务和开展名词审定工作的网络平台。

一个网站或者软件系统的设计开发,前期的需求分析与架构设计非常关键。在明确了网站的性质和功能定位之后,下一步就是要明确数据库和网站的设计与开发原则,这是方案设计和选型的前奏。在《天文学名词》数据库新版网站的设计和实现中,确定了如下的原则:

(1) 服务型网站:虽然本网站是名词委的官方网站,但根本目的是为广大用户提供天文学名词相关的服务。栏目设置和界面设计都要从用户的使用需求出发。用户需要怎样的信息和功能,就提供怎样的信息和功能。

(2) 简单实用:数据库技术和网络信息技术日新月异,但新版系统仍将延续原有系统的简约风格,内容第一、形式第二。同时,适当采用新的技术为用户带来更好的用户体验。

(3) 无需注册,开放式服务:这既简化了系统的设计和开发难度,又大大方便用户的使用。新网站旨在为用户提供更好的名词服务,重在寻求社会效益。

2 需求分析

自2002年对外开放以来,《天文学名词》数据库很快赢得了广大科研人员和公众的认可,用户群不断扩大。用户在使用过程中反馈给天文学名词委许多良好的建议和需求。归纳起来,原有系统存在的问题集中表现在以下几个方面。

(1) 数据库结构过于简单,信息量严重不足。原有数据库中仅提供了外文词条和中文译名两项

内容。许多非常重要的信息，比如释义、范例、台湾译名等，都需要补充，从而让用户能更准确和方便地学习和使用。

(2) 网站与用户互动的功能很少，可参与性不强。作为天文学名词委名词审定工作的网络平台和服务平台还需要增加更多的功能，比如允许用户对词条进行评论，为词条补充信息，对用户提交的评论和推荐的新词进行便捷的管理等功能。名词审定是一件非常严肃但同时又极其耗费精力的工作，充分利用用户和社区的力量是必然的选择。维基百科^[6]和国内爱好者自发推出的维基天文^[7]都是很好的例子。

(3) 缺乏网站主办方天文学名词委的资料和信息。作为天文学名词委的官方网站，需要为用户提供全面、准确的背景资料和相关信息。

找到了现有系统的缺点和努力的方向，下一步的工作就是把这些需要的功能和应该提供的信息落实到用户界面上。用户通过界面上的菜单、按钮和链接等要素完成对系统的使用。

3 用户界面设计

图 1 是《天文学名词》数据库新版网站的导航菜单布局。本系统定位很明确：名词服务、名词审定。因此，新版网站仅设计了 4 个主栏目，即“首页”、“天文词典”、“名词审定”和“消息动态”。为了突出网站的服务作用，“首页”上仅提供 3 个关键要素：词库搜索表单、词库词条统计和随机展示的词条，如图 2。

“天文词典”是为用户使用数据库专门开辟的区域，其中的各项子菜单都与数据库浏览和查询检索有关，具体包括：

(1) 词典检索：和“首页”类似的一个词典检索入口；

(2) 分类浏览：按照词条的类别属性对常用的词条，比如星座、节气等，提供快速浏览；

(3) 新词推荐：用户推荐新词条的表单；

(4) 加入讨论：为用户介绍如何通过邮件列表加入针对名词的讨论；

(5) 词典下载：根据天文学名词委开放使用名词数据库的决定，允许用户下载数据库中全部词条的基本信息；

(6) 工具 & 服务：一些小工具和服务的集锦，比如网页嵌入代码、RSS 种子、中文输入法软件用户词典等；

(7) 常检词条：对用户已提交数据库来进行检索的词语进行统计排名，供更多用户参考；

(8) 常见问题：对用户在使用数据库过程中经常遇到的问题进行解答。

“词条详情”、“词条评论”和“词义补充”都是针对特定词条而言的，所以这些功能设置在了词典查询结果的显示页面中，为每个词条动态的生成相应的链接，而没有出现在图 1 的菜单列表中。如图 3，在显示词条详情的页面中不但提供了词义补充和评论的链接，还设置了到 Google、维基百科、维基天文等网站的快速通道。

“名词审定”主栏目提供了关于天文学名词委全面的介绍和相关背景资料，以及名词审定工作的网络平台。这些内容体现出该网站作为天文学名词委官方网站的身份。名词审定工作的功能放在了一个需要密码才能访问的内部“工作区”。

“消息动态”是本网站的新闻消息区域。在其下的“名词动态”中为用户提供了国内外与科技名词，特别是天文学名词相关的一些消息。“网站更新”则记录了本数据库网站系统的历史变更。



图 1 网站导航菜单

Fig. 1 Website navigation menu

这样的栏目设置把为用户提供服务放在了首要位置,同时也充分表达了其作为天文学名词委官方网站的含义。



图2 《天文学名词》数据库网站首页
Fig. 2 Home page of the website “Glossary of Astronomical Terms”



图3 词条详情页面中提供的丰富链接
Fig. 3 Functional links provided in a webpage of dictionary entry details

4 数据库设计

用户界面是底层信息的前台展示窗口。为了前台能够顺畅、高效地从后台提取所需的信息,后台数据库的设计至关重要。尽管《天文学名词》数据库网站是一个相对简单的系统,为了能完整地支持用户界面上设计的各种功能,在后台设计了一个由7个表构成的数据库系统。各表结构及其相互关系如图4。这7个表的名称和用途分别是:

- (1) 主词表(astrodict): 存储名词的基本信息,是本数据库的核心数据表;
- (2) 用户评论表(comments): 存储用户针对主词表中的词条提交的评论和词义补充信息;
- (3) 新词推荐表(recommendation): 存储用户推荐的新词信息;
- (4) 词条状态表(status): 保存主词表中词条的状态类别;
- (5) 词条来源表(source): 保存主词表中词条的出处信息;
- (6) 词库使用情况表(usage): 记录用户每次检索数据库的时间和所输入的检索词等信息;
- (7) 词库文本下载情况表(download): 记录用户下载数据库全文的时间等信息。

下面以主词表为例来详细说明它的结构设计及其与其他表的联系。主词表字段结构如表1。表中,词条唯一ID标识(UID)为主键(PK),通过它来标识每个词条记录。为了提高检索效率,对需要进行检索的字段都添加索引。词条状态(status)和词条来源(source)为外键(FK),真实文本保存在词条状态表(status)和词条来源表(source)中。系统在输出检索结果时需要对这3个表进行联合查询。用户评论表(comments)中的词条唯一ID标识(UID)是一个外键,指向主词表中的UID字段。这样的设计保证了用户只能对现有词库中的词条进行评论和补充词义。其余3个表则独立工作,相互之间没有约束关系。

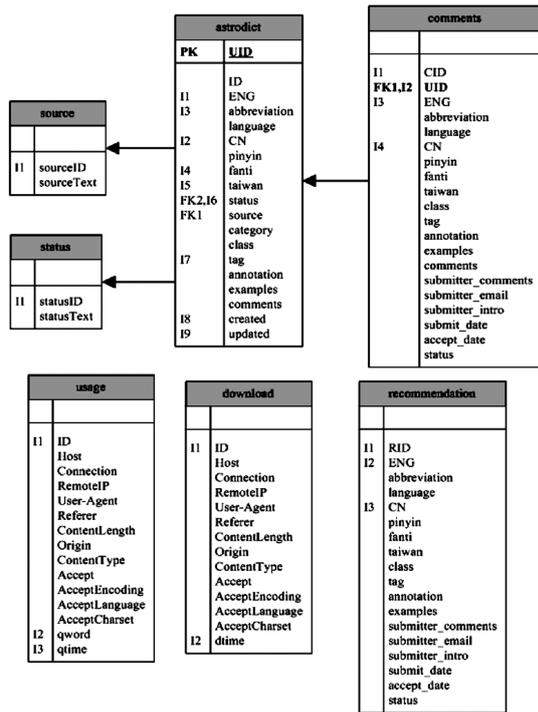


图 4 数据库结构

Fig. 4 Database structure

表 1 主词表(astrodict)字段结构

Table 1 Structure of the fields in the main table "astrodict"

字段	类型	字段含义
ID	整型	自动增长的 ID
UID	整型	词条唯一 ID 标识
ENG	文本	外文名
abbreviation	文本	缩写
language	文本	语种
CN	文本	中文名
pinyin	文本	拼音
fanti	文本	中文繁体
taiwan	文本	台湾名
status	文本	词条状态
source	文本	词条来源
category	文本	学科分类
class	文本	内容分类
tag	文本	分类标签
annotation	文本	释义
examples	文本	范例
comments	文本	备注
created	日期	收录时间
updated	日期	最近更新时间

5 开发与实现

对于一个网络应用系统而言,程序开发的工作量往往只是全部工作量的一小部分。应用系统的规模越大越是如此。在大型应用系统的设计开发中,前期需求分析与系统设计还有后期的调试与测试是最需要投入的地方。

《天文学名词》数据库网站是一个小型的网络应用系统,程序开发与实现采用了有“建站黄金组合”之称的 LAMP 方案^[8],即“Linux 操作系统”+“ApacheWWW 引擎”+“MySQL 开源数据库”+“PHP 网络程序设计语言”。Linux 操作系统是天文学家最常用的工作平台。目前主流的 Linux 发行版本都集成了对 Apache、MySQL 和 PHP 的支持,所以采用 LAMP 组合对《天文学名词》这样的网站开发来说是最佳的选择。

为了便于维护,对不同网页间的共同部分,比如抬头、页角、菜单等都定义了模版,不同网页间变化的只是各自特有的部分。图 5 给出了网站页面的 PHP 模版。

为了统一页面风格,同时方便页面布局的调整,全面采用层叠式样式表(CSS)控制显示输出。

为了保证用户评论、新词推荐等内容的有效性,杜绝恶意用户的破坏,防范网络蜘蛛等程序的垃圾信息提交,在兼顾用户使用方便的前提下,在词条评论和新词推荐的表单页面中设置了有效的保护机制。

新版网站 2010 年 4 月 18 日正式上线。4 月 19 日至 5 月 18 日一个月内,数据库的查询次数是 6 607 次。上线前的一个月,3 月 19 日至 4 月 18 日,原有系统的检索次数为 4 029 次。访问量增加了 64%。同时新版上线以来已经收到推荐新词 22 条,补充词义和评论 3 条。此外,还收到了数十封来自用户的非常正面的电子邮件反馈。运行情况表明新版的设计目标已经实现。

```
<?php
require("../global.inc"); //加载全局参数和常量
include("../dict_header.php"); //加载网页抬头部分
include("../menulist.php"); //加载网页菜单部分
?>
<!-- 具体网页的内容 -->
Web Page Content here

<?php
include("../dict_footer.php"); //加载网页页角部分
?>
```

图5 网页公用模版

Fig. 5 Common template of webpages for the database

6 讨论与展望

天文学名词委的基本职能是开展天文学名词的研究和审定工作,促进天文学的国际交流。《天文学名词》数据库是天文学名词委数十年来工作成果的具体体现,也是《天文学名词》数据库网站的核心财富。

内容是核心,界面是形式,技术是连接内容和形式的纽带。《天文学名词》数据库网站自从2002年开放运行以来,不断丰富内容和服务形式,从最初简单的名词数据库演变成了科技人员和广大公众使用天文学名词、参与名词工作、学习名词知识,天文学名词委开展名词审定工作的综合性平台。

随着天文学研究和信息技术的快速发展,《天文学名词》数据库网站和天文学名词委名词审定工作都面临着巨大的机遇和挑战。最大的挑战和困难是,新词的不断涌现及其审定工作所需的大量智力投入与天文学名词委人力和时间严重短缺的矛盾。虽然Web 2.0、维基百科、社区网络等技术能够在一定程度上让互联网用户和社区力量参与到名词审定的工作中,但科技名词审定工作的严肃性决定了这种参与只能是有限程度的,无法从根本上解决问题。

同时我们也看到,《天文学名词》的发展空间是巨大的。近年来迅速发展的以虚拟天文台^[9-10]为典型代表的天文信息技术,力图实现全球范围内天文学研究和教育等资源的大融合。《天文学名词》数据库作为一种高度浓缩的信息库,如果能与国内外的天文数据库、文献库等各种信息资源互联互通并最终完成融合,必将是一种质的飞越。当然,其中涉及很多技术挑战和大量的工作。如果实施,也必将加剧前面提到的挑战和矛盾。

虽然这种理想目标不可能一步实现,但可以从多个小的方面入手,逐步向这个大的方向前进,比如:在《天文学名词》和SIMBAD^[11]天体证认数据库以及其他语种的天文学名词数据库之间建立基本的程序接口;以《天文学名词》数据库为核心开发输入法用户词典、增值应用程序等。这些设想都已经列入了《天文学名词》数据库的未来规划中。

致谢:《天文学名词》数据库网站能取得现在的业绩,是历届天文学名词审定委员会委员,两岸三地乃至海外的同仁,关心和支持天文学名词工作的各界朋友共同努力的结果。国家天文台、国家大科学工程“郭守敬望远镜(LAMOST)”为本系统提供了运行环境,中国虚拟天文台(China-VO)为本系统提供了技术支持。感谢中国天文学会和全国科学技术名词审定委员会。

参考文献:

- [1] 沈明惠. 中国天文学会成立天文学名词审定委员会[J]. 天文学进展, 1983, 1(1):120.

- Shen Minghui. The Establishment of the Astronomical Terminology Lommittee of CAS [J]. *Progress in Astronomy*, 1983, 1(1): 120.
- [2] 全国科学技术名词审定委员会 [EB/OL]. <http://www.cnctst.gov.cn/>.
- [3] 李启斌. 对当前名词审定工作的几点建议[J]. *科技术语研究*, 2000, 2(3): 30–31.
- [4] 崔辰州. 充分利用现代信息技术开展科技新词工作[J]. *中国科技术语*, 2010(3): 19–21.
- [5] 《天文学名词》数据库网站[EB/OL]. <http://www.lamost.org/astrodict/>.
- [6] 维基百科[EB/OL]. <http://en.wikipedia.org/>.
- [7] 维基天文[EB/OL]. <http://wiki.skylook.org/>.
- [8] LAMP[EB/OL]. http://en.wikipedia.org/wiki/LAMP_%28software_bundle%29.
- [9] Cui Chenzhou, Zhao Yongheng. Worldwide R&D of Virtual Observatory [C]//W J Jin, I Platais, M A C Perryman. *A Giant Step: from Milli- to Micro-arcsecond Astrometry: Proceedings of the International Astronomical Union(2007)2008*, 3(Symposium S248): 563–564.
- [10] 崔辰州, 赵永恒. 中国虚拟天文台研发策略与重点 [J]. *天文研究与技术——国家天文台台刊*, 2004, 1(3): 203–209.
- Cui Chenzhou, Zhao Yongheng. Research and Development Strategy and Focuses of Chinese Virtual Observatory [J]. *Astronomical Research & Technology——Publications of National Astronomical Observatories of China*, 2004, 1(3): 203–209.
- [11] SIMBAD astronomical object database[EB/OL]. <http://simbad.u-strasbg.fr/simbad/>.

Design and Implementation of a New Version of the Website “Glossary of Astronomical Terms”

Cui Chenzhou¹, Yu Heng², Bian Yulin³

(1. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China, Email: ccz@bao.ac.cn;

2. Department of Astronomy, Beijing Normal University, Beijing 100875, China;

3. Shanghai Scientific & Technological Education Press, Shanghai 200235, China)

Abstract: The website “Glossary of Astronomical Terms” is an online platform for the Astronomical Terminology Committee of the Chinese Astronomical Society (ATCC) to officially approve astronomical terms and to provide public query/search or related services of these terms. With the progress of ATCC’s approving work and increasing requirements from users on the functions of the database, the original website is no longer sufficient for these functions and information amounts. We design and develop a new version of the website with mature database and network technologies. The new version covers the entire process of establishing astronomical terminology, i. e. term collection, approving, release, query, and use. The new version satisfies requirements from users and the ATCC. We describe in the paper design principles and implementation methods of the new version. Following an analysis of the requirements, we emphatically give the details of the design and implementation of the backstage database and user interface. The website “Glossary of Astronomical Terms” is a typical simple database-driven network application system. Work and experiences as described in the paper provide practical references for similar systems.

Key words: Glossary of Astronomy Terms; Requirement analysis; Database design; Website design; Implementation