LAMOST
科学数据库开发与进展

VO-enabled LAMOST

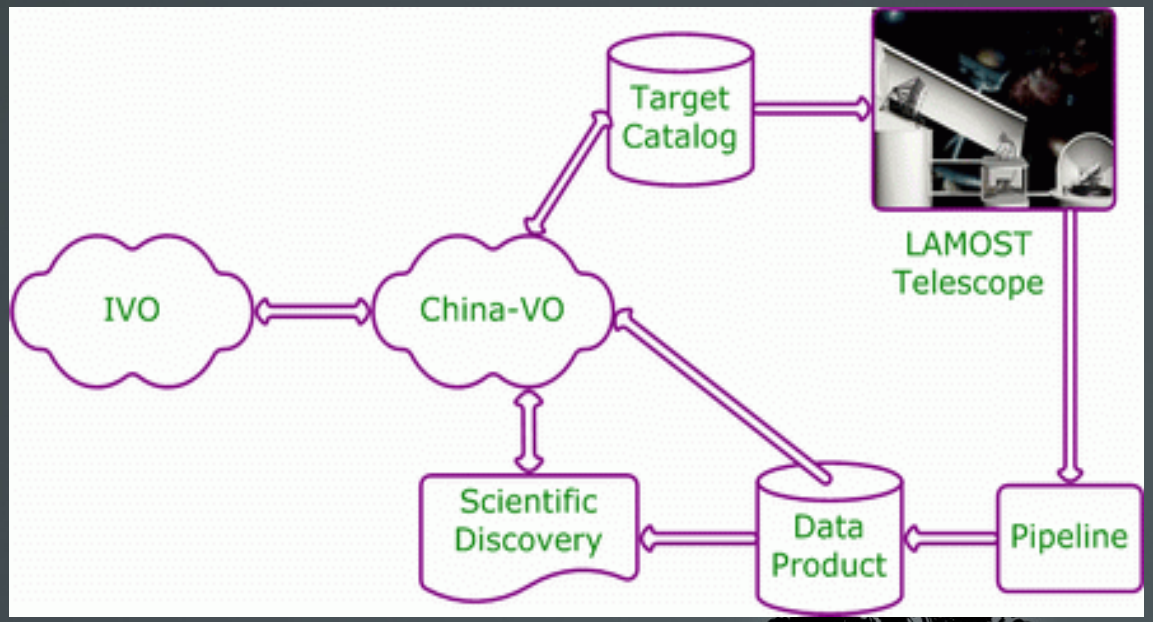
何勃亮
hebl@lamost.org

2008-11-28

国家天文台
LAMOST

目录

- LAMOST 科学数据库
- LAMOST 科学数据库与 Virtual Observation
- 项目开发进展
- 未来设想
- Q & A



LAMOST 科学数据库简介

- 数据库简介
- 系统功能模块
- 数据库开发应用环境
- 管理的数据
- 数据库 数据表
- 可以进行的操作
- 数据存储
- Web 页面功能设计
- 其它

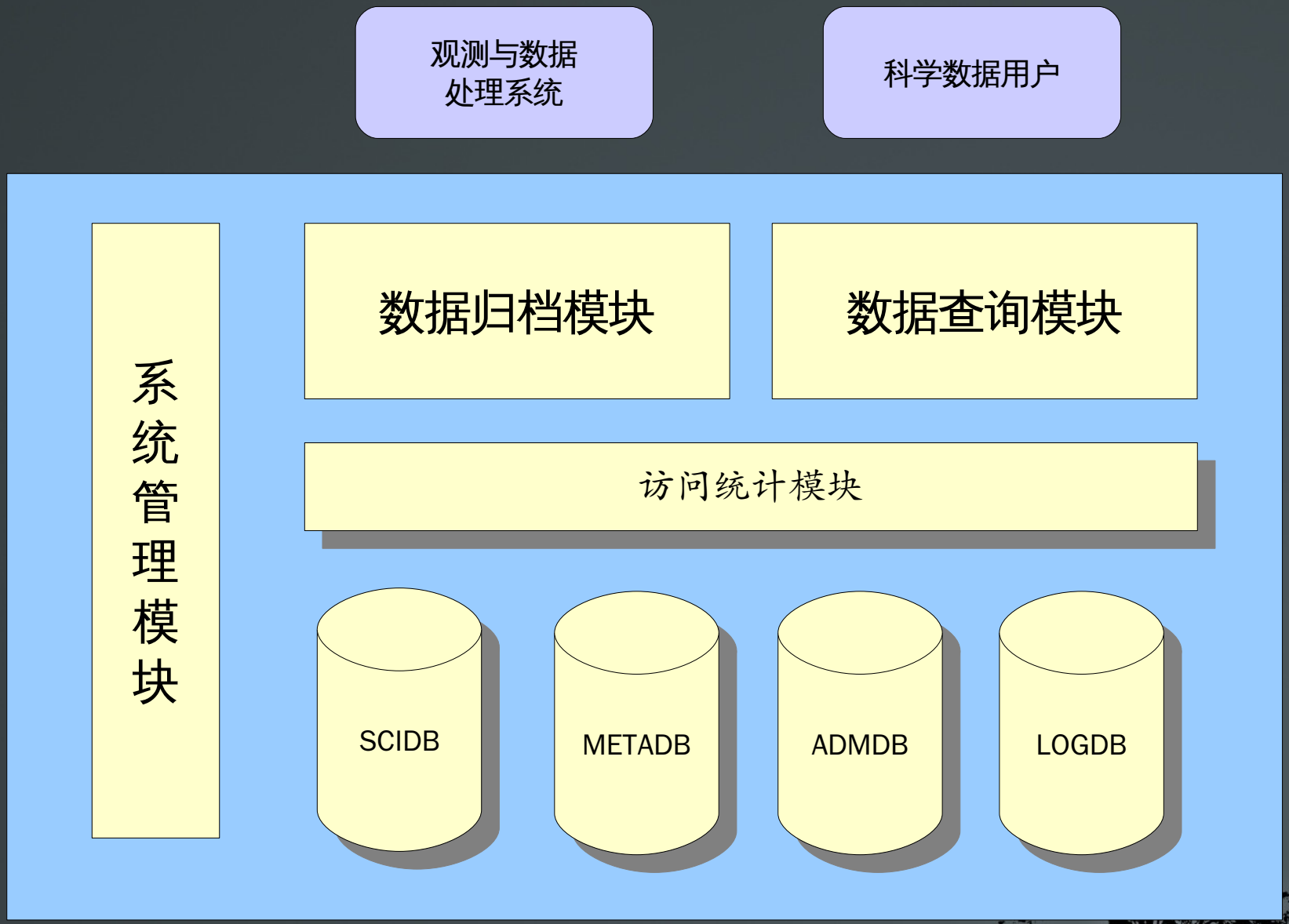
LAMOST 的运行观测

- LAMOST
  - 4000 根光纤
  - 4000 / 250 => 16 光谱仪
  - Red & blue 16 x 2 => 32 CCD 照相机
  - CCD 相机 => 4k x 4k pixels => 16 bits color deep
- 每个观测天区 (plate)，3 次暴光，每次暴光 30 分钟，每晚观测 7.5~9 小时，大约 5 个 plate
- 每晚观测数据： 4K x 4K x 2 x 32 x 3 x 5= 15.36 (GB)
- 每晚所有的数据： 20 GB
- 每年： 20 x 200= 4,000 GB per year
- 5 年周期： 4 x 5 year = 20 TB

数据库简介

- 数据库是大型天文望远镜数据管理的重要方式。
- LAMOST 科学数据库是 LAMOST 数据管理子系统的主要组成部分，它负责的是对观测数据（原始数据、输入星表等）进行管理和维护，数据库将对 LAMOST 需要用到的数据和观测的数据进行高效的运行管理。
- 数据库的用户群：管理用户（系统管理、数据维护）、数据用户（巡天观测用户、项目内部科学家、公开用户）。
- 数据产品
  - 2-D multi-fiber spectrum images (pravite)
  - 1-D spectra (to the public 18 months later)
  - Catalogs (to the public 24 months later)

系统功能模块


观测与数据
处理系统
科学数据用户
系统管理模块
数据归档模块
数据查询模块
访问统计模块
SCIDB
METADB
ADMDB
LOGDB

与其它子系统的接口要求

- SSS、OCS、TCS 及其他各观测相关子系统需要尽可能完整的保留自己收集的数据，并把这些数据以 XML 文件、ASCII 文件或者 MySQL 数据库等格式保存起来，并定期向数据管理系统归档。数据管理子系统决定最终归档的数据格式和目录格式。
- 望远镜观测获取的原始数据以 FITS 文件的形式提交数据管理系统进行归档。数据管理系统进行归档的同时进行必要的备份。
- 数据处理系统，包括二维数据处理系统和一维数据处理系统，从数据管理系统获取 Pipeline 的输入数据。数据处理系统向数据管理系统提交的处理结果包括包含完整信息的 FITS 文件和包括完整信息的星表。
- 各个相关子系统需要采用统一的科学数据库模型，通过一套统一的 ID 系统进行操作。

软件开发应用环境

- 服务环境：Linux（RHEL、Fedora、etc.）
- 数据库软件：MySQL、PostgreSQL、Oracle 等
- 开发语言：PHP、Javascript、Perl、C、C++、Java 等
- 应用技术：AJAX、CGI、Web Service 等
- 开发的服务：B/S、Web Service 等
- 数据库中间层：抽象对数据库的操作以适应数据库可能的迁移

管理的数据

- 观测数据
  - 每次观测的数据记录
  - 观测的一维二维光谱数据
  - 天光信息等
- 巡天星表
  - 二维光谱星表
  - 一维光谱星表
- 工作星表
- 系统运行数据（各个系统运行日志、环境气象信息、观测信息等）
- 其他需要保存的数据

数据库 数据表

LAMOST 科学数据库包含的数据表分为四类：

- Sci_* 科学数据类表
  - 工作星表
  - 运行信息
  - 观测信息
- Meta_* 元数据表
  - sci_* 数据表的元数据
  - 其它元数据
- Admin_* 管理类数据表
  - 用户信息
  - 管理员信息
  - 用户空间管理
- Log_* 日志数据表

数据表

- ID System
- Global unique targetID
  - sourceID: source catalog ID, for “SDSS” it is “SDSS”, for “GSC 2.3” it is “GSC23”, etc.
  - sourceTargetID: the original ID for a source in the source catalog
  - targetID: a unique ID for each object of LAMOST observation. (HTM)
  - plateID: telescope pointing ID, i.e. sky area ID. (HTM)
  - runID: exposure ID
  - obsID: observation ID, an observation comprises one plate and multiple exposure
  - spectrographID: spectrograph ID
  - chipID: CCD chip ID
  - fiberID: fiber ID
- Indexing: HTM

可以进行的操作

- 搜索查询类操作
  - 观测结果的查询
  - 工作星表查询
  - 观测运行信息查询
  - 数据导出（HTML、FITS、VOTable、CSV、其它）
- 数据结果的操作
  - 星表在线查看 （VOPlot、Topcat等在线离线使用）
  - 光谱的在线查看（VOSpec、Specview、Aladin等的在线离线使用）
  - 其它
- 数据管理

数据存储

- iRods (https://www.irods.org)
- VOSpace
- 同步、备份等

Web 页面 功能设计

- 首页
  - 用户登录
  - 管理员登录
- 用户空间界面
  - 用户空间
  - 用户进行数据查询
  - 用户星表
  - 查询结果存储、打包、下载

Web 页面 功能设计（续）

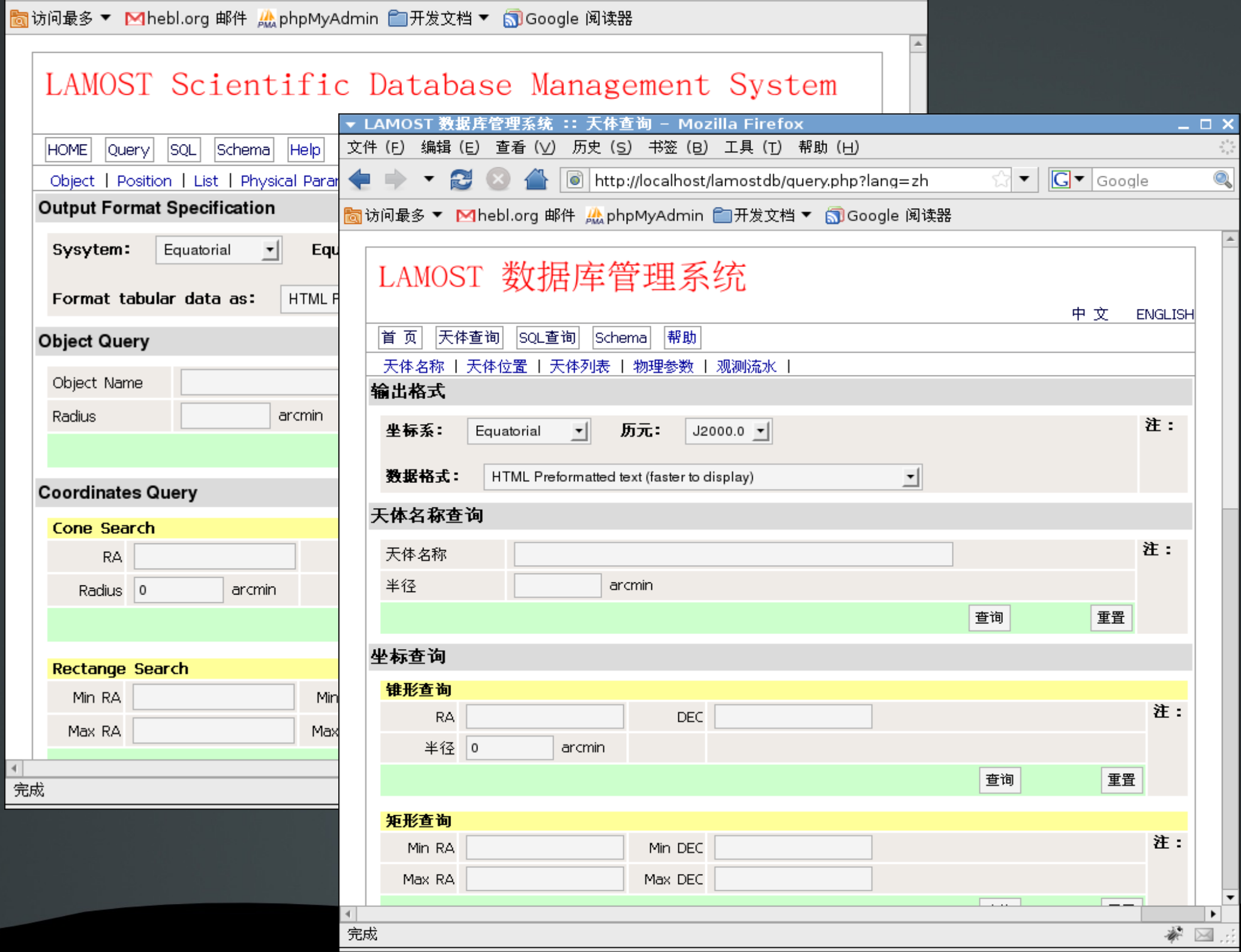
- 管理界面
  - 数据库查看管理
  - 数据导入导出备份等
  - 文件压缩归档
  - 数据检查
  - 用户管理
    - 角色管理
    - 授权管理
    - 监控、预警
    - 其它
  - 访问日志管理
    - 访问统计
    - 查询类型统计
    - 下载统计
    - 系统信息
  - 版本管理等

Web 页面 功能设计（续）

- 查询首页
  - 根据天体名称查询（ Simbda ）
  - 根据天体位置查询（锥形、矩形等）
  - 根据观测流水查询（观测时间、观测仪器等）
  - 根据物理参数查询
  - 根据证认信息查询
  - 根据数据质量查询（望远镜运行状况、环境天气状况、数据处理状况等）
  - 运行日志查询
  - 直接 SQL 语句查询
  - 其它查询

Web 页面 功能设计（续）

- 查询结果显示
  - Html 表格输出
  - FITS 表格 / 图像输出
  - VOTable 表格 / 图像输出
  - CSV 输出
  - 查询结果打包下载
- VOSpace 与 iRods
- 协同工作环境集成
- 支持中英文的自动切换


访问最多
hebl.org 邮件
phpMyAdmin
开发文档
Google 阅读器
LAMOST Scientific Database Management System
HOME
Query
SQL
Schema
Help
Object | Position | List | Physical Parar
Output Format Specification
Sysytem:
Equatorial
Format tabular data as:
Object Query
Object Name
Radius
arcmin
Coordinates Query
Cone Search
RA
Radius
0
arcmin
Rectange Search
Min RA
Max RA
完成
LAMOST 数据库管理系统 :: 天体查询 - Mozilla Firefox
文件 (F) 编辑 (E) 查看 (V) 历史 (S) 书签 (B) 工具 (T) 帮助 (H)
http://localhost/lamostdb/query.php?lang=zh
Google
LAMOST 数据库管理系统
中 文 ENGLISH
首 页
天体查询
SQL查询
Schema
帮助
天体名称 | 天体位置 | 天体列表 | 物理参数 | 观测流水 |
输出格式
坐标系:
Equatorial
历元:
J2000.0
注：
数据格式:
HTML Preformatted text (faster to display)
天体名称查询
天体名称
半径
arcmin
查询
重置
坐标查询
锥形查询
RA
DEC
半径
0
arcmin
查询
重置
矩形查询
Min RA
Min DEC
Max RA
Max DEC
完成

LAMOST 科学数据库与 VO

VO-enabled LAMOST，VO 时代的 LAMOST，在 LAMOST 科学数据库的开发和应用中将对 VO 协议和 VO 应用进行充分的支持。

在 LAMOST 科学数据库中，将处处见 VO，处处使用 VO

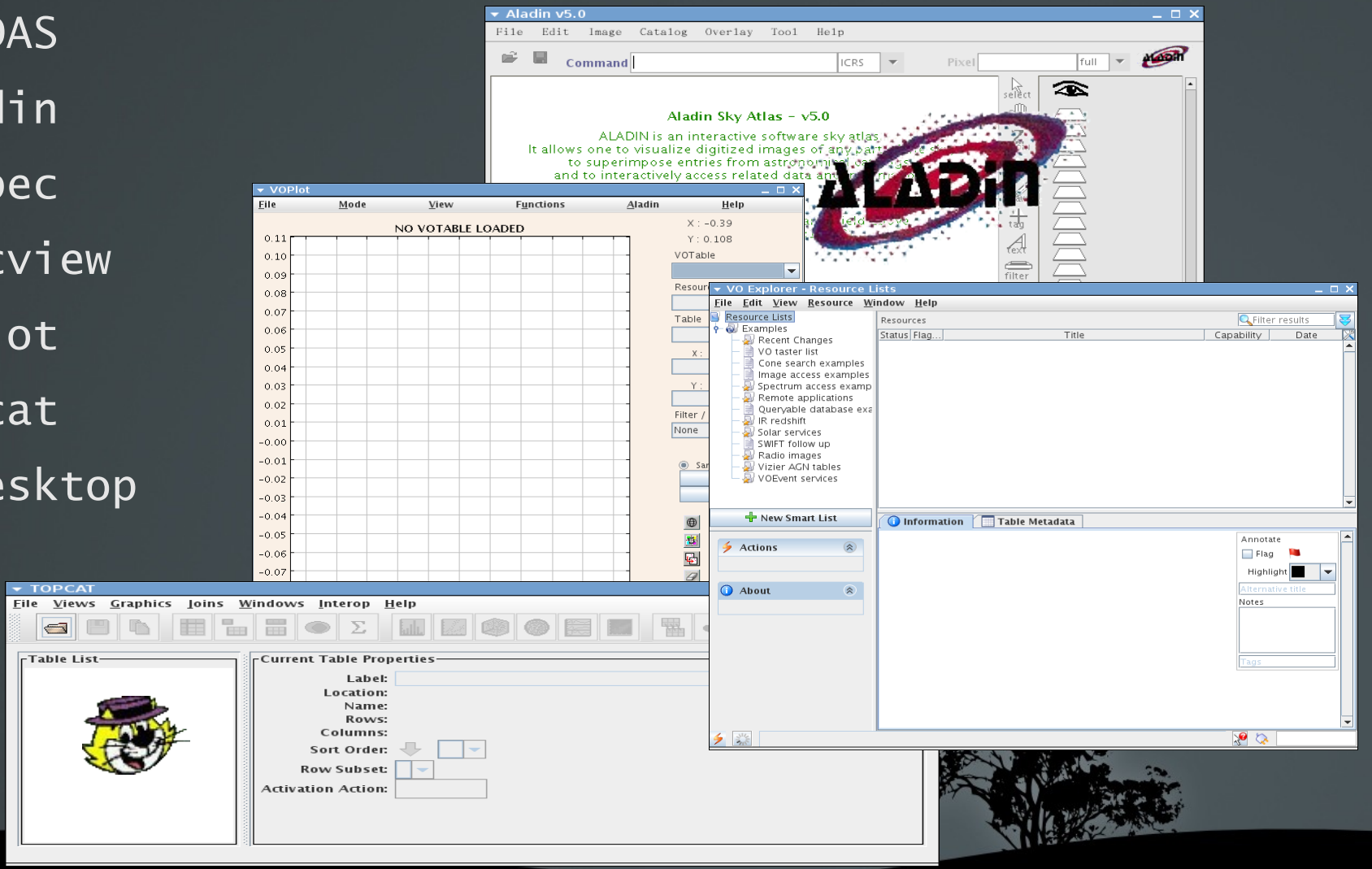
支持的 VO 协议

- VOTable
- Table Access Protocol (TAP)
- Metadata/UCD
- Simple Cone Search
- Simple Image Access (SIA)
- Simple Spectral Access (SSA)
- Space-Time Coordinate Metadata
- VO Regiestry
- VOSpace
- Footprint
- ADQL
- ...

http://www.ivoa.net/Documents/

支持的 VO 应用

- VO 应用的支持（大部分支持在线的应用）
  - VO-DAS
  - Aladin
  - VOSpec
  - Specview
  - VOPlot
  - Topcat
  - VODesktop
  - 其它
-



LAMOST 科学数据库与 VO （示例）

VO 协议和 VO 应用的支持是通过网络实现的，下面是几个实现的简单方式。用户或者 VO 应用都可以通过 HTTP 的 GET 方式获得数据，一般的，默认查询返回的是 HTML，也可以根据需要返回 FITS 或 VOTable 等。

- Simple Cone Search
  - cones.php?ra=2.2&dec=2.3&sr=2&output=votable
  - rects.php?ra1=1&dec1=2&ra2=3&dec2=4&output=html
- Simple Image Access
  - sia.php?pos=2.2,2.3&size=0.012
- Simple Spectrum Access
  - ssa.php?version=1.0&pos=2.2,2.3&band=J&size=0.2
- ....

项目开发进展

- 2008 年开始开发，已完成部分功能性设计
- 2009 年 1 月完成第一个 Beta 版（基本的数据管理功能，可以配合 LAMOST 观测进行使用）
- 2009 年 5 月完成第一版的定型，可以供 LAMOST 日常观测使用

未来设想

- 协同工作环境的集成
- 更加人性化设计
- 针对用户的需求继续进行改进
- ...

謝謝！

Q & A